(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 November 2002 (28.11.2002)

PCT

(10) International Publication Number
WO 02/094865 A1

(51) International Patent Classification[7]: C07K 14/00, C12N 5/10, 15/11, 15/12, 15/62, 15/63

(21) International Application Number: PCT/US02/16445

(22) International Filing Date: 24 May 2002 (24.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/293,380 : 24 May 2001 (24.05.2001) US

(71) Applicant *(for all designated States except US)*: **SMITHKLINE BEECHAM CORPORATION** [US/US]; One Franklin Plaza, Philadelphia, PA 19101 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KLIEWER, Steven, Anthony** [US/US]; 3453 Potomac Avenue, Dallas, TX 75205 (US). **MAGLICH, Jodi, Marie** [US/US]; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US). **MOORE, John, Tomlin** [US/US]; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US). **MOORE, Linda, Becker** [US/US]; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US). **WILLSON, Timothy, Mark** [GB/US]; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US).

(74) Agents: **LEVY, David, J.** et al.; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/094865 A1

(54) Title: NONHUMAN PREGNANE X RECEPTOR SEQUENCES FOR USE IN COMPARATIVE PHARMACOLOGY

(57) Abstract: Polynucleotides and polypeptides of canine, primate, porcine and fish pregnane X receptor (PXR), as well as expression vectors and host cells for expression of these PXR receptors, are provided. Also provided are methods for screening for modulators of these PXR receptors and using these receptors for comparative pharmacology and in selection of appropriate preclinical animal models predictive of human PXR activity.

NONHUMAN PREGNANE X RECEPTOR SEQUENCES FOR USE IN COMPARATIVE PHARMACOLOGY

Field of the Invention

The present invention relates to nonhuman pregnane X receptor (PXR) polynucleotide and polypeptide sequences isolated from fish, canines, porcine and primates and their use in comparative pharmacology.

Background of the Invention

Members of the cytochrome P450 family of hemoproteins are critical in the oxidative metabolism of a wide variety of endogenous substances and xenobiotics, including various carcinogens and toxins (Nebert et al. (1987) Ann. Rev. Biochem. 56:945-993). In man, P450 3A4 monooxygenase, also referred to as CYP3A4 monooxygenase, plays a major role in the biotransformation of drugs due to its abundance in liver and intestine and its broad substrate specificity. P450 3A4 catalyzes the metabolism of >60% of all drugs that are in use including steroids, immunosuppressive agents imidazole antimycotics, and macrolide antibiotics (Maurel, P. in Cytochrome P450: metabolic and toxicological aspects (ed. Ioannides, C.) 241-270 (CRC Press, Inc. Boca Raton, FL 1996).

Expression of P450 3A4 is induced both *in vivo* and in primary hepatocytes in response to treatment with a variety of compounds including, but not limited to, commonly used drugs such as the glucocorticoid dexamethasone, the antibiotic rifampicin, the antimycotic clotrimazole, and the hypocholesterolemic agent lovestatin (Maurel, P. in Cytochrome P450: metabolic and toxicological aspects (ed. Ioannides, C.) 241-270 (CRC Press, Inc. Boca Raton, FL 1996); Guzelian, P.S. in Microsomes and Drug Oxidation (eds. Miners, J.O., Birkett, D.J., Drew, R. & McManus, M.) 148-155 (Taylor and Francis, London, 1988)). The inducibility of P450 3A4 expression

levels coupled with its broad substrate specificity represent the basis for many drug interactions in patients undergoing combination therapy. Thus, analysis of the effects of new compounds on P450 3A4 gene expression is an important aspect in drug development.

While attempts have been made to develop *in vivo* and *in vitro* assays with which to profile the effects of compounds on P450 3A4 expression levels, poorly understood species-specific variations have limited the utility of using animals and their tissues for testing purposes. Thus, analysis of the effects of new compounds on P450 3A4 gene expression has been largely restricted to laborious assays involving human liver tissue.

Nuclear hormone receptors comprise a large superfamily of ligand-modulated transcription factors that, in part, mediate responses to steroids, retinoids, and thyroid hormones (for review see Beato *et al.*, (1995) *Cell* 83:851-857; Kastner *et al.*, (1995) *Cell* 83:859-869; Mangelsdorf and Evans, (1995) *Cell* 83:841-850). Detailed analysis of the receptors, most notably the steroid class of receptors, has revealed multiple discrete functional modules within the family that display generalized functional characteristics (Tzukerman *et al.*, (1994) *Mol. Endocrinol.* 8:21-30). A variable amino-terminal domain (A/B) is present that typically contains a strong and autonomous activation function (AF1), shown to be critical for cell and target gene specificity (Tora *et al.*, (1988) *Nature* 333: 677-684). A more carboxyl-terminal central region contains a DNA binding domain (DBD) characterized by two C4-type zinc fingers. The DBD binds to specific genomic response elements and thereby regulates the transcriptional activity of select genes containing the response elements. At the distal carboxyl terminus, a ligand binding domain (LBD) is present containing a highly conserved second transactivation function (AF2) that is important for hormone-dependent transcriptional transactivation (Lanz and Rusconi, (1994) *Endocrinology* 135:

2183-2195). Sequences that function in nuclear localization, receptor dimerization, and interaction with heat-shock proteins (Gronemeyer and Laudet, (1995) *CCQ* 2:1173-1308) are also present within the nuclear receptor substructure. Through the coordinated action of these separate functional domains, nuclear receptor activation by ligand culminates in modulation of target gene expression (Tsai and O'Malley, (1994) *Ann. Rev. Biochem.* 63: 451-486) and in certain cases, cross-talks with other cell signaling pathways such as the NF-kB (Stein and Yang, (1995) *Mol. Cell. Biol.* 15: 4971-4979) and AP-1 (Paech *et al.*, (1998) *Science* 277: 1508-1510). Ultimately, ligand alters nuclear receptor function by altering the constellation of protein-protein interactions in which the receptor is engaged (for review, see Freedman, (1999) *Cell* 97: 5-8). The molecular details underlying the multitude of cellular effects mediated by nuclear receptors are the subject of intense research activity.

For the nuclear receptor for pregnane X, referred to herein as PXR (Unified Nomenclature Committee designation NR1I2), it has been shown that PXR cell-based and binding assays are predictive of *in vivo* effects on the cytochrome P450 3A4 gene (Kliewer et al. (1998) Cell 92, 73-8273-82; Lehmann et al. (1998) J. Clin. Invest. 102, 1016-1023; Jones et al. (2000) Mol. Endo. 27-39).

WO 99/48915 discloses human PXR which binds to the CYP promoter rifampicin/dexamethasone response element in cytochrome P450 3A4. Also disclosed are nucleic acid sequences encoding human PXR, as well as vectors and host cells for expression of the human receptor, and methods for using this receptor *in vitro* to screen compounds for their ability to modulate P450 3A4 expression in humans.

However, nonhuman animal PXRs, particularly animals well accepted for use in preclinical studies, would also be useful in the development *in vitro* and *in vivo* animal models for profiling the effects of compounds on P450 3A4 expression

levels and to select preclinical models predictive of effects in humans. Additionally, dissecting the broader biological physiological role of PXR beyond cytochrome P450 gene induction is facilitated by an understanding of which compounds activate PXR in animal models of interest. Thus, comparative pharmacology of PXR opens the possibility of extending the utility of PXR in drug development and toxicity assays as well as validating this receptor as a useful target in disease.

Mouse PXR has been cloned and sequenced (Kliewer et al. (1998) Cell 92:73-82). Rat and rabbit PXR have also been cloned and sequenced (Jones et al. (2000) Molecular Endocrinology 14(1):27-39; Zhang et al. (1999) Archives of Biochemistry and Biophysics 368(1):14-22).

The present invention relates to PXRs for other nonhuman animals.

Summary of the Invention

Polynucleotide and polypeptide sequences for pregnane X receptors (PXR) isolated from fish, canines, porcine and primates are provided. The novel PXR sequences are useful as screening targets for the identification and development of selective PXR compounds. These agents are particularly useful in PXR comparative pharmacology and selecting appropriate animal models for preclinical studies predictive of effects in humans.

Accordingly, the present invention provides isolated PXR polypeptides comprising:

(a) an amino acid sequence of SEQ ID NO: 2, 4, 6 or 8;

(b) a variant of an amino acid sequence as defined in (a) which modulates P450 3A4 levels or activity; or

(c) a fragment of (a) or (b) which modulates P450 3A4 levels or activity.

According to another aspect of the invention there is provided polynucleotides encoding polypeptides of the

invention, said polynucleotides comprising:

(a) a nucleic acid sequence of SEQ ID NO: 1, 3, 5, or 7;

(b) a nucleic acid sequence which hybridizes under stringent conditions to the nucleic acid sequence as defined in (a);

(c) a nucleic acid sequence that is degenerate as a result of the genetic code to the nucleic acid sequences as defined in (a) or (b); or

(d) a nucleic acid sequence having at least 60% identity to the nucleic acid sequences as defined in (a), (b) or (c).

Other related aspects of the present invention include expression vectors comprising polynucleotides of the invention which are capable of expressing a polypeptide of the invention, host cells comprising an expression vector of the invention, methods of producing a polypeptide of the invention which comprise maintaining a host cell of the invention under conditions suitable for obtaining expression of the polypeptide and isolating said polypeptide, antibodies specific for a polypeptide of the invention, and transgenic nonhuman animals expressing a mutant PXR or a PXR from another species.

The present invention also provides methods for identification of substances that modulate PXR activity and/or expression. In one embodiment, the method comprises contacting a polypeptide, polynucleotide, expression vector or host cell of the invention with a test substance and determining the effect of the test substance on the activity and/or expression of the polypeptide to determine whether the test substance modulates PXR activity and/or expression. In another embodiment, the test substance is administered to a nonhuman transgenic animal expressing a mutant PXR or a PXR from a different species and the effects of the test substance on expression and/or activity of the receptor are examined.

In addition, the present invention relates to compounds which modulate PXR activity and which are identifiable by the methods referred to above.

Brief Description of the Sequences

SEQ ID NO: 1 shows a nucleotide and amino acid sequence of a ligand binding domain of a canine pregnane X receptor.

SEQ ID NO: 2 shows the amino acid sequence of the ligand binding domain of the canine pregnane X receptor as depicted in SEQ ID NO:1.

SEQ ID NO:3 shows a nucleotide and amino acid sequence of a primate pregnane X receptor.

SEQ ID NO:4 shows the amino acid sequence of the primate pregnane X receptor as depicted in SEQ ID NO:3.

SEQ ID NO:5 shows a nucleotide and amino acid sequence of a ligand binding domain of a porcine pregnane X receptor.

SEQ ID NO:6 shows the amino acid sequence of the ligand binding domain of the porcine pregnane X receptor as depicted in SEQ ID NO:5.

SEQ ID NO:7 shows a nucleotide and amino acid sequence of a ligand binding domain of a Zebrafish pregnane X receptor.

SEQ ID NO:8 shows the amino acid sequence of the ligand binding domain of the Zebrafish pregnane X receptor as depicted in SEQ ID NO:7.

Brief Description of the Figures

Figure 1 is a bargraph showing the activation of a primate PXR in the presence of various steroids and xenobiotics.

Figure 2 is a bargraph showing the activation of a canine PXR in the presence of various steroids and xenobiotics.

Figure 3 is a bargraph showing the activation of a porcine PXR in the presence of various steroids and xenobiotics.

Figure 4 is a bargraph showing the activation of a fish PXR in the presence of various steroids and xenobiotics.

Figure 5 is a bargraph showing the activation of a primate PXR in the presence of various bile acids.

Figure 6 is a bargraph showing the activation of a canine PXR in the presence of various bile acids.

Figure 7 is a bargraph showing the activation of a porcine PXR in the presence of various bile acids.

Figure 8 is a bargraph showing the activation of a fish PXR in the presence of various bile acids.

Detailed Description of the Invention

Throughout the present specification and the accompanying claims the words "comprise" and "include" and variations such as "comprises", "comprising", "includes" and "including" are to be interpreted inclusively. That is, these words are intended to convey the possible inclusion of other elements or integers not specifically recited, where the context allows.

The present invention relates to nonhuman animal orthologs of pregnane X receptors, referred to herein as PXRs, and variants thereof. More specifically, the present invention relates to PXRs isolated from a canine, porcine, primate or fish. Nucleotide sequence information for the full length monkey PXR of the present invention is provided in SEQ ID NO: 3. A polypeptide sequence of the monkey PXR is also provided in SEQ ID NO: 4. Sequence information for the ligand binding domains of the dog, pig and Zebrafish PXRs of the present invention is provided in SEQ ID NO: 1, 5 and 7 (nucleotide and amino acid), respectively. Polypeptides of the ligand binding domains of the dog, pig and Zebrafish PXRs are also provided in SEQ ID NO: 2, 6, and 8.

Polypeptides of the invention consist essentially of the amino acid sequences of SEQ ID NO: 2, 4, 6 or 8, a variant of that sequence, or a fragment of either thereof. Polypeptides of the invention may be in a substantially isolated form. It

will be understood that the polypeptide may be mixed with carriers or diluents which will not interfere with the intended purpose of the polypeptide and still be regarded as substantially isolated. A polypeptide of the invention may also be in a substantially purified form, in which case it will generally comprise the polypeptide in a preparation in which more than 50%, e.g. more than 80%, 90%, 95% or 99%, by weight of the polypeptide in the preparation is a polypeptide of the invention. Routine methods, can be employed to purify and/or synthesize the polypeptides according to the invention. Such methods are well understood by persons skilled in the art, and include techniques such as those disclosed in Sambrook *et al*, Molecular Cloning: a Laboratory Manual, 2nd Edition, CSH Laboratory Press, 1989, the disclosure of which is included herein in its entirety by way of reference.

The term "variant" refers to a polypeptide that has a same essential character or basic biological functionality as the selected PXR. As demonstrated herein, biological functionalities of nonhuman PXRs vary between species. By "selected PXR receptor", as used herein, it is meant a canine, primate, porcine or fish PXR as described herein. Further, for purposes of this invention a variant polypeptide is preferably one which binds to the same ligand as one or more of the nonhuman animal PXRs described herein. Preferably the polypeptide modulates P450 3A4 expression in primates, canines, porcine and/or fish. A polypeptide having a same essential character as a selected PXR of the present invention can be identified by monitoring for activation of the selected PXR by an inducer of P450 3A4. A full-length variant polypeptide is preferably one which includes the entire ligand binding domain of the selected PXR.

In another aspect of the invention, a variant is one which does not show the same activity as the selected PXR but rather inhibits a basic function of the PXR. For example, a variant polypeptide is one which inhibits activation of a

selected PXR upon exposure to an inducer of P450 3A4, for example by binding to a PXR ligand to prevent activity mediated by ligand binding to the selected PXR.

Typically, polypeptides with more than about 65% identity preferably at least 80% , at least 85% or at least 90% and particularly preferably at least 95%, at least 97% or at least 99% identity, with the amino acid sequences of SEQ ID NO: 2, 4, 6 or 8 are considered as variants of the proteins. Identity can be determined using a program such as a BLAST sequence alignment program. Such variants may include allelic variants and the deletion, modification or addition of single amino acids or groups of amino acids within the protein sequence, as long as the peptide maintains a basic biological functionality of the selected PXR.

Amino acid substitutions may be made, for example from 1, 2 or 3 to 10, 15, 20 or 30 substitutions. The modified polypeptide generally retains the same activity as a selected PXR. Conservative substitutions may be made, for example according to the following Table 1. Amino acids in the same block in the second column and preferably in the same line in the third column may be substituted for each other.

TABLE 1

ALIPHATIC	**Non-polar**	G A P
		I L V
	Polar-uncharged	C S T M
		N Q
	Polar-charged	D E
		K R
AROMATIC		H F W Y

Exemplary substitutions are shown in Table 2.

TABLE 2

Original Residue	Exemplary Substitutions
Ala	Gly; Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Ala
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg
Met	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Shorter polypeptide sequences, also referred to herein as "fragments" are within the scope of the invention. For example, a peptide of at least 15, 20 or 30 amino acids or up to 50, 60, 70, 80, 100, 150 or 200 amino acids in length is considered to fall within the scope of the invention as long

as it demonstrates a basic biological functionality of a selected PXR. In particular, but not exclusively, this aspect of the invention encompasses the situation when the protein is a fragment of the complete protein sequence and may represent particularly, the A/B domain (e.g., about amino acids 1-40 of SEQ ID NO:4), the DBD (e.g., about amino acids 41-105 of SEQ ID NO:4) or the LBD (e.g., about amino acids 106-434 of SEQ ID NO:4 or SEQ ID NO:2, 6 or 8), alone or in combination. Such fragments can be used to construct chimeric receptors preferably with another nuclear receptor, more preferably with another member of the family of nuclear receptors involved in P450 3A4 regulation such as the CAR nuclear receptor, or as intermediates in the production of the full length sequences.

Such fragments of PXR or a variant thereof can also be used to raise anti-PXR antibodies. In this embodiment, the fragment may comprise an epitope of a selected PXR polypeptide and may otherwise not demonstrate the ligand binding or other properties of the selected PXR.

Polypeptides of the invention may be chemically modified, e.g. post-translationally modified. For example, they may be glycosylated or comprise modified amino acid residues. They may also be modified by the addition of histidine residues to assist their purification or by the addition of a signal sequence to promote insertion into the cell membrane. Such modified polypeptides fall within the scope of the term "polypeptide" of the invention.

The invention also includes nucleotide sequences that encode for canine, primate, porcine or fish PXR or a variant thereof, as well as nucleotide sequences which are complementary thereto. The nucleotide sequence may be RNA or DNA including genomic DNA, synthetic DNA or cDNA. Preferably

the nucleotide sequence is a DNA sequence, and most preferably a cDNA sequence. Nucleotide sequence information for the canine, primate, porcine and Zebrafish PXRs of the present invention are provided in SEQ ID NO: 1, 3, 5 and 7, respectively. Such nucleotides can be isolated from cells of the selected species, namely, canine, primate, porcine or Zebrafish, or synthesized according to methods well known in the art, as described by way of example in Sambrook *et al*, 1989.

Typically a polynucleotide of the invention comprises a contiguous sequence of nucleotides which is capable of hybridizing under selective conditions to the coding sequence or the complement of the coding sequence of either SEQ ID NO: 1, 3, 5 or 7.

The polynucleotide can be in single-, double- or triple-stranded form. In addition, polynucleotide, as used herein, can refer to triple-stranded regions comprising RNA or DNA or both. The strands in such regions may be from the same molecule or from different molecules. As used herein the term polynucleotide includes nucleic acids that contain one or more modified (e.g., tritylated) or unusual (e.g., inosine) bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are polynucleotides as that term is intended herein.

A polynucleotide of the invention can hybridize to the coding sequence or the complement of the coding sequence of SEQ ID NO: 1, 3, 5 or 7 at a level significantly above background. Background hybridization may occur, for example, because of other cDNAs present in a cDNA library. The signal level generated by the interaction between a polynucleotide of the invention and the coding sequence or complement of the

coding sequence of SEQ ID NO: 1, 3, 5 or 7 is typically at least 10-fold, preferably at least 100-fold, as intense as interactions between other polynucleotides and the coding sequence of SEQ ID NO: 1, 3, 5 or 7. The intensity of interaction may be measured, for example, by radiolabeling the probe, e.g. with ^{32}P. Selective hybridization may typically be achieved using conditions of medium to high stringency. However, such hybridization may be carried out under any suitable conditions known in the art (see Sambrook *et al*, 1989). For example, if high stringency is required suitable conditions include from 0.1 to 0.2 x SSC at 60°C up to 65°C. If lower stringency is required suitable conditions include 2 x SSC at 60°C.

The coding sequence of SEQ ID NO: 1, 3, 5 or 7 may be modified by nucleotide substitutions, for example from 1, 2 or 3 to 10, 15, 25, 50 or 100 substitutions. The polynucleotide of SEQ ID NO: 1, 3, 5 or 7 may alternatively or additionally be modified by one or more insertions and/or deletions and/or by an extension at either or both ends. A polynucleotide may include one or more introns. For example, in one embodiment, the polynucleotide may comprise genomic DNA. Additional sequences such as signal sequences which may assist in insertion of the polypeptide in a cell membrane may also be included. The modified polynucleotide generally encodes a polypeptide which has the same activity as a selected PXR. Alternatively, a polynucleotide encodes a ligand-binding portion of a polypeptide or a polypeptide which inhibits an activity of a selected PXR. Degenerate substitutions may be made and/or substitutions may be made which would result in a conservative amino acid substitution when the modified sequence is translated, for example as shown in Table 1 or 2

above.

A nucleotide sequence which is capable of selectively hybridizing to the complement of the DNA coding sequence of SEQ ID NO: 1, 3, 5, or 7 will generally have at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 98% or at least 99% sequence identity to the coding sequence of SEQ ID NO: 1, 3, 5 or 7 over a region of at least 20, preferably at least 30, for instance at least 40, at least 60, more preferably at least 100 contiguous nucleotides, or most preferably over the full length of SEQ ID NO: 1, 3, 5 or 7.

For example the UWGCG Package provides the BESTFIT program which can be used to calculate homology (for example used on its default settings) (Devereux *et al.*,(1984) *Nucleic Acids Res.* 12: 387-395). The PILEUP and BLAST algorithms can be used to calculate homology or line up sequences (typically on their default settings), for example as described in Altschul (1993) *J. Mol. Evol.* 36: 290-300; Altschul *et al* (1990) *J. Mol. Biol.* 215: 403-410.

Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.* 1990). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extensions for the word

hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1992) *Proc. Natl. Acad. Sci.* USA 89: 10915-10919) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm performs a statistical analysis of the similarity between two sequences; see e.g., Karlin and Altschul (1993) *Proc. Natl. Acad. Sci.* USA 90: 5873-5787. One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a sequence is considered similar to another sequence if the smallest sum probability in comparison of the first sequence to the second sequence is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Any combination of the above mentioned degrees of sequence identity and minimum sizes may be used to define polynucleotides of the invention, with the more stringent combinations (i.e. higher sequence identity over longer lengths) being preferred. Thus, for example a polynucleotide which has at least 90% sequence identity over 25, preferably over 30 nucleotides forms one aspect of the invention, as does a polynucleotide which has at least 95% sequence identity over

40 nucleotides.

The nucleotides according to the invention have utility in production of the polypeptides according to the invention. Produciton may take place *in vitro, in vivo* or *ex vivo.* The nucleotides may be involved in recombinant protein synthesis, as therapeutic agents in their own right, utilized in gene therapy techniques, and/or utilized in the production of nonhuman transgenic animals. Nucleotides complementary to those encoding PXR, or antisense sequences, may also be used therapeutically.

Polynucleotides of the invention may be used as primers, e.g. PCR primers or primers for an alternative amplification reaction, or as probes, e.g. polynucleotides detectably labeled by conventional means using radioactive or non-radioactive labels. In addition, the polynucleotides may be cloned into vectors.

Such primers, probes and other fragments will preferably be at least 10, preferably at least 15 or at least 20, for example at least 25, at least 30 or at least 40 nucleotides in length. They will typically be up to 40, 50, 60, 70, 100 or 150 nucleotides in length. Probes and fragments can be longer than 150 nucleotides in length, for example up to 200, 300, 400, 500, 600, 700 nucleotides in length, or even up to a few nucleotides, such as five or ten nucleotides, short of the coding sequence of SEQ ID NO: 1, 3, 5 or 7.

The polynucleotides of the present invention are also useful in the production of chimeric receptors or fusion proteins having a PXR component which comprises at least a DNA binding domain or a ligand binding domain of a canine, primate, porcine or fish PXR and a non-PXR derived sequence.

Non-PXR derived sequences can be selected so as to be suitable for the purpose to be served by the chimeric receptor. Examples of such sequences include, but are not limited to, glutathione-S-transferase, the DNA binding domain of yeast transcription factor GAL4 and other DNA binding domains such as the DNA binding domains for estrogen or glucocorticoid receptors, and the viral VP16 transcriptional activation domain. Chimeric receptors of the present invention may further comprise a detectable label such as a radioactive or fluorescent label. The chimeric receptors may also be bound to a solid support such as glass or plastic particles or plates or a filter.

The present invention also includes expression vectors that comprise nucleotide sequences encoding the polypeptides or variants thereof of the invention. Such expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may be necessary, and which are positioned in the correct orientation, in order to allow for protein expression. Other suitable vectors would be apparent to persons skilled in the art based upon teachings provided herein and what is known in the art. By way of further example in this regard we refer to Sambrook *et al.* 1989.

Polynucleotides according to the invention may also be inserted into the vectors described above in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense polynucleotides may also be produced by synthetic means. Such antisense polynucleotides may be used as test compounds in the

assays of the invention or may be useful therapeutically.

Preferably, a polynucleotide of the invention used in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. For example, a regulatory sequence, such as a promoter, "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequence.

The vectors may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example, for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell. The vectors may also be adapted to be used *in vivo*, for example in a method of gene therapy or in the production of nonhuman transgenic animals, preferably mice. Additional vector components known in the art are suitable for use in the vectors of the present invention and include, for example, processing sites such as a polyadenylation signal, ribosome binding sites, RNA splice sites, and transcriptional termination sequences.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which

expression is designed. Examples of yeast promoters which can be used in the present invention include *S. cerevisiae* GAL4 and ADH promoters, *and S. pombe* NMT1 and ADH promoters. Viral promoters can also be used. Examples of viral promoters include, but are not limited to, the Moloney murine leukemia virus long terminal repeat (MMLV LTR), the Rous sarcoma virus (RSV) LTR promoter, the SV40 promoter, the human cytomegalovirus (CMV) IE promoter, adenovirus, HSV promoters (such as the HSV IE promoters), or HPV promoters, particularly the HPV upstream regulatory region (URR). A mammalian promoter useful in the present invention is the metallothionein promoter which can be induced in response to heavy metals such as cadmium and β-actin promoters. Tissue-specific promoters are especially preferred. All these promoters, as well as additional promoters useful in the present invention, are readily available in the art.

The vector may further include sequences flanking the polynucleotide giving rise to polynucleotides which comprise sequences homologous to eukaryotic genomic sequences, preferably mammalian genomic sequences or viral genomic sequences. This allows the introduction of the polynucleotides of the invention into the genome of eukaryotic cells or viruses by homologous recombination. In particular, a plasmid vector comprising the expression cassette flanked by viral sequences can be used to prepare a viral vector suitable for delivering the polynucleotides of the invention to a mammalian cell. Other examples of suitable viral vectors include herpes simplex viral vectors and retroviruses, including lentiviruses, adenoviruses, adeno-associated viruses and HPV viruses. Gene transfer techniques using these viruses are known to those skilled in the art. Retrovirus vectors for

example may be used to stably integrate the polynucleotide into the host genome. Replication-defective adenovirus vectors by contrast remain episomal and therefore allow transient expression.

The invention also includes cells that have been modified to express the PXR polypeptide or a variant thereof. Such cells include transient, or preferably stable higher eukaryotic cell lines such as mammalian cells or insect cells, lower eukaryotic cells such as yeast, or prokaryotic cells such as bacterial cells. Examples of cells which may be modified by insertion of vectors encoding for a polypeptide according to the invention include, but are not limited to, mammalian HEK293T, CHO, HeLa and COS cells.

A polypeptide of the invention may also be expressed in cells of a transgenic non-human animal, preferably a mouse. Accordingly, transgenic non-human animals expressing a PXR polypeptide of the invention are also included within the scope of the invention. For example, transgenic mice can be generated that express a selected PXR of the present invention as well as the endogenous mouse PXR gene. Mice can also be generated in which the endogenous PXR gene is knocked out and then replaced by the selected PXR polynucleotide of the present invention. Transgenic animals can also be generated that express isoforms of a selected PXR as well as mutant alleles of the PXRs of the present invention. Transgenic animals developed by these methods can be used to screen compounds for drug interactions and toxicities and to study the regulation of P450 3A4 *in vivo*.

According to another aspect, the present invention also relates to antibodies, specific for a polypeptide of the invention. Such antibodies are for example useful in

purification, isolation or screening methods involving immunoprecipitation techniques or, indeed, as therapeutic agents in their own right.

Antibodies can be raised against specific epitopes of the polypeptides according to the invention. Such antibodies may be used to block ligand binding to the receptor. An antibody, or other compound, "specifically binds" to a protein when it binds with preferential or high affinity to the protein or polypeptide for which it is specific but does not substantially bind or binds with only low affinity to other proteins. A variety of protocols for competitive binding or immunoradiometric assays to determine the specific binding capability of an antibody are well known in the art (see for example Maddox *et al*, (1993) *J. Exp. Med.* 158: 1211-1226). Such immunoassays typically involve the formation of complexes between the specific protein and its antibody and the measurement of complex formation.

Antibodies of the invention may be antibodies to the canine, primate, porcine or fish polypeptides or fragments thereof. For the purposes of this invention, the term "antibody", unless specified to the contrary, includes fragments which bind a polypeptide of the invention. Such fragments include Fv, F(ab') and $F(ab')_2$ fragments, as well as single chain antibodies. Furthermore, the antibodies and fragments thereof may be chimeric antibodies, CDR-grafted antibodies or humanized antibodies.

Antibodies may be used in methods for detecting polypeptides of the invention in a biological sample. In these methods, an antibody of the invention is first provided. A biological sample is then incubated with the antibody under conditions which allow for the formation of a complex between

the antibody and the polypeptide or antigen and the amount of antibody-polypeptide complex formed is determined. Various methods for determining formation of an antibody-antigen complex are well known to those of skill in the art.

For purposes of the present invention, by "biological sample" it is meant to include, but is not limited to, tissue extracts, blood, serum, saliva, urine, cerebral spinal fluid, and bile.

Antibodies of the invention may be bound to a solid support and/or packaged into kits in a suitable container along with suitable reagents, controls, instructions, etc. Antibodies may be linked to a revealing label and thus may be suitable for use in methods of *in vivo* PXR imaging.

Antibodies of the invention can be produced by any suitable method. Means for preparing and characterizing antibodies are well known in the art, see for example Harlow and Lane (1988) "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. For example, an antibody may be produced by raising antibody in a host animal against the whole polypeptide or a fragment thereof, for example an antigenic epitope thereof, herein after the "immunogen".

A method for producing a polyclonal antibody comprises immunizing a suitable host animal, for example an experimental animal, with the immunogen and isolating immunoglobulins from the animal's serum. In this method, the animal is inoculated with the immunogen. Blood is subsequently removed from the animal and the IgG fraction purified.

A method for producing a monoclonal antibody comprises immortalizing cells which produce the desired antibody.

Hybridoma cells can be produced by fusing spleen cells from an inoculated experimental animal with tumor cells (Kohler and Milstein (1975) *Nature* 256: 495-497). An immortalized cell producing the desired antibody may be selected by a conventional procedure. Hybridomas are then grown in culture or injected intraperitoneally for formation of ascites fluid or into the blood stream of an allogenic host or immunocompromised host.

For the production of both monoclonal and polyclonal antibodies, the experimental animal is suitably a goat, rabbit, rat or mouse. If desired, the immunogen may be administered as a conjugate in which the immunogen is coupled, for example via a side chain of one of the amino acid residues, to a suitable carrier. The carrier molecule is typically a physiologically acceptable carrier. The antibody obtained may be isolated and, if desired, purified.

A further aspect of the present invention relates to *in vitro* (cell-free) and *in vivo* (cell-based) assays that can be used to profile the effects of compounds, particularly potential new drugs, on P450 3A4 levels in various species. These assays can take any of a variety of forms. Since compounds that activate PXR function are inducers of P450 3A4 gene expression, binding and activation assays using selected PXRs of the present invention provide efficient means to identify compounds expected to activate P450 3A4 in a selected species.

Binding assays of the present invention include cell free assays in which a selected PXR, or the ligand binding domain of a selected PXR (alone or present as a fusion protein), is incubated with a test compound which, advantageously, bears a detectable label (e.g. a radioactive or fluorescent label).

The selected PXR, or ligand binding domain thereof, free or bound to the test compound, is then separated from free test compound using any of variety of techniques (e.g., using gel filtration chromatography (for example, on Sephadex G50 spin columns) or through capture on a hydroxyapatite resin). The amount of test compound bound to the selected PXR or ligand binding domain thereof, is then determined via detection of the label.

An alternative approach for detecting radiolabeled test compound bound to a selected PXR, or ligand binding domain thereof, is a scintillation proximity assay (SPA). In this assay, a bead (or other particle) is impregnated with scintillant and coated with a molecule that can capture the selected PXR, or ligand binding domain thereof (e.g., streptavidin-coated beads can be used to capture biotinylated PXR ligand binding domain). Radioactive counts are detected only when the complex of radiolabeled test compound and the selected PXR, or ligand binding domain thereof, is captured on the surface of the SPA bead bringing the radioactive label into sufficient proximity to the scintillant to emit a signal. This approach has the advantage of not requiring the separation of free test compound from bound (Nichols et al, Anal. Biochem. 257:112-119 (1998)).

Assays to determine whether a test compound interacts with a selected PXR ligand binding domain can also be performed via a competition binding assay. In this assay, the selected PXR, or ligand binding domain thereof, is incubated with a compound known to interact with the selected PXR, which compound, advantageously, bears a detectable label (e.g., a radioactive or fluorescent label). A test compound is added to the reaction and assayed for its ability to compete with

the labeled compound for binding to the selected PXR, or ligand binding domain thereof. A standard assay format employing a step to separate free known (labeled) compound from bound, or an SPA format, can be used to assess the ability of the test compound to compete.

To determine if a test compound activates a selected PXR, and thus induces P450 3A4 expression, the ligand binding domain of the selected PXR is prepared (e.g., expressed) as a fusion protein (e.g., with glutathione-S-transferase (GST), a histidine tag or a maltose binding protein). The fusion protein and coactivator (either or both advantageously labeled with a detectable label, e.g., a radiolabel or fluorescent tag) are incubated in the presence and absence of the test compound and the extent of binding of the coactivator to the fusion protein determined. The induction of interaction in the presence of the test compound is indicative of an activator of the selected PXR.

PXR activation assays in accordance with the invention can be carried out using a full length PXR and a reporter system comprising one or more copies of the DNA binding site recognized by the PXR binding domain. More preferably, however, the activation assays are conducted using established chimeric receptor systems. For example, the ligand binding domain of a selected PXR can be fused to the DNA binding domain of, for example, yeast transcription factor GAL4, or that of the estrogen or glucocorticoid receptor. An expression vector for the chimera (e.g.,a GAL4-PXR chimera) can be transfected into host cells (e.g., CV-1, HuH7, HepG2 or Caco2 cells) together with a reported construct. The reporter construct may comprise one or more (e.g., 5) copies of the DNA binding site recognized by the binding domain present in the

chimera (e.g., the GAL4 DNA binding site) driving expression of a reporter gene (e.g., CAT, SPAP or luciferase). Cells containing the constructs are then treated with either vehicle alone or vehicle containing test compound, and the level of expression of the reporter gene determined. In accordance with this assay, enhancement of expression of the reporter gene in the presence of the test compound indicates that the test compound activates the selected PXR and thus can function as an inducer of CYP3A4 gene expression in that species.

Another format suitable for use in connection with the present invention is the yeast two-hybrid assay. This is an established approach to detect protein-protein interactions that is performed in yeast. Protein #1, representing the bait, is expressed in yeast as a chimera with a DNA binding domain (e.g., GAL4). Protein #2, representing the predator, is expressed in the same yeast cell as a chimera with a strong transcriptional activation domain. The interaction of bait and predator results in the activation of a reporter gene (e.g., luciferase or β-galactosidase) or the regulation of a selectable marker (e.g., LEU2 gene). This approach can be used as a screen to detect, for example, ligand-dependent interactions between a selected PXR and other proteins such as coactivator proteins (e.g., SRCI, TIFI, TIF2, ACTR) or fragments thereof (Fields et al., Nature 340:245-2.46 (1989)).

Still another format is the ligand-induced complex formation (LIC) assay. This assay detects ligand-mediated effects on nuclear receptor-DNA interactions. A selected PXR (or, minimally, DNA and/or ligand binding domains thereof) can be incubated with its heterodimeric partner RXR in the presence of DNA representing an established PXR/RXR binding site. Test-compounds can be assayed for their ability to

either enhance or interfere with binding of the PXR/RXR heterodimer to DNA (Forman et al, Proc. Natl. Acad. Sci. USA 94:4312-4317 (1997).

In a preferred embodiment, the screening assay takes the form of a FRET (Fluorescence Resonance Emission Transfer assay (Nichols *et al.* (1998) *Anal. Biochem.* 257:112-119). This screening assay comprises the steps of exposing a sample portion comprising the donor located at a first position and the acceptor located at the second position to light at a first wavelength capable of inducing a first electronic transition in the donor. The donor comprises a complex of a lanthanide chelate and a lanthanide capable of binding the chelate. The spectral overlap of the donor emission and acceptor absorption is sufficient to enable energy transfer from the donor to the acceptor as measured by a detectable increase in acceptor luminescence. In a preferred embodiment, a SRC-1 (LCD2, 677-696) lanthanide chelate is used. Preferably, the lanthanide element comprises Europium and the signal chelate comprises Europium bound to a PXR of the present invention. A signal pair comprising Europium bound to the PXR and APC (allophycocyanin) bound to SRC-1 (see, e.g., Parks *et al.* (1999) *Science* 284:1365-1368) can also be used.

Suitable test compounds which can be screened in the above assays include combinatorial libraries, defined chemical entities and compounds, peptide and peptide mimetics, oligonucleotides, natural product libraries such as display libraries (e.g. phage display libraries), and antibody products.

Typically, organic molecules, preferably small organic molecules which have a molecular weight of from 50 to 2500 daltons, are screened. Candidate test compounds can be

biomolecules including, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs and combinations thereof. Such test compounds are obtained from a wide variety of sources including libraries of synthetic and natural compounds. Further, known pharmacological agents can be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs.

Test compounds can be used in an initial screen of, for example, 10 compounds per reaction, and the compounds of these batches which show inhibition or activation re-screened individually. Test compounds may be screened at a concentration of from 1 nM to 1000 μM, preferably from 1 μM to 100 μM, more preferably from 1 μM to 10 μM. Preferably, the activity of a test compound is compared to the activity shown by a known activator or inhibitor. A test compound which acts as an inhibitor preferably produces a 50% inhibition of activity of the receptor. Alternatively a test compound which acts as an activator preferably produces 50% of the maximal activity produced using a known activator.

Comparative pharmacology involves the use of a nonhuman animal model to either predict the effects of a compound in humans, or provide a contrast to the effects of a compound in humans. Accordingly, results from assays such as described above with the PXRs of the present invention provide important information with respect to whether a compound of interest modulates the receptor similarly to or differently from the analogous human receptor.

Further, comparison of activation of human PXRs with activation of a non-human PXR of the present invention or PXRs from other nonhuman animals which are used as models in

preclinical studies is useful in selection of preclinical animal models predictive of affects of a test compound on P450 3A4 in humans. In this method, *in vitro* activation of human PXR in the presence of a test compound is compared with *in vitro* activation of PXRs from various preclinical animal models, including, but not limited to the PXRs of the present invention, in the presence of the same test compound. A PXR from a preclinical animal model exhibiting similar *in vitro* activation to the human PXR in the presence of the test compound is indicative of the preclinical animal model being predictive of the affects of the test compound on P450 3A4 in humans.

Activation of the nonhuman PXRs of the present invention was examined in the presence of various steroids, xenobiotics, and bile acids. Steroids and xenobiotics screened for effects on activation of the nonhuman PXRs of the present invention included: pregnenolone 16a-carbonitrile (PCN); rifampicine; 3 amino ethyl benzoate; TCPOBOP (1,4-bis[2-(3,5-dichloropyridyloxy)]benzene); epoxycholesterol; RU 486; omeprazole; primidone; ethosuximide; nifedipine; metyrapone; reserpine; trans-nanochlor; androstanol; clofibrate; clofibric acid; troglitazone; 6, 16-dimethylpregnenolone; pregnenolone; 17a-OH pregnenolone; progesterone; 17a-OH-progesterone 5b-pregnane 3,20-dione; corticosterone; cortisone; DHEA (dehydroepiandrosterone); DHT (dihydrotestosterone); spironolactone; b-estradiol; tamoxifen; dexamethasone; dex-t-butylacetate; hydrocortisone; d-aldosterone; cyproterone acetate; hyperforin; phenobarbital; carbamazepine; phenytoin; clotrimazole; SR12813; lovastatin; mevastatin; squalestatin; and chlorpromazine. Bile acids screened for effects on activation of the nonhuman PXRs of the present invention

included: 12-ketolithocholic acid; 3,6-diketocholanic acid; 3,7-diketocholanic acid; 3a, 7a-dihydroxy-12-ketocholanic acid; 6-ketolithocholic acid; 7,12-diketolithocholic acid; 7-ketodeoxycholic acid; 7-ketolithocholic acid; chenodeoxycholic acid; cholic acid; dehydrolithocholic acid; deoxycholic acid; 7-ketodeoxycholic acid methyl ester; glychochenodeoxycholic acid; glycocholic acid; glycodehydrocholic acid; glycodeoxycholic acid; glycohyocholic acid; glychohyodeoxycholic acid; taurodeoxycholic acid; glycolithocholic acid; hyocholic acid; hyodeoxycholic acid; lithocholic acid; murocholic acid; taurochenodeoxycholic acid; taurocholanic acid; taurocholic acid; taurodehydrocholic acid; taurohyocholic acid; taurodeoxycholic acid; taurolithocholic acid; tauro-b-muricholic acid; ursocholanic acid; ursodeoxycholic acid; a-muricholic acid; b-muricholic acid; 5b-cholanic acid-7a,12a-diol-3-one; and 5b-cholanic acid-3,7,12-trione. The co-transactivation assay described in Example 2 was used to screen these compounds. Activation was assessed via measurement of levels of secreted placental alkaline phosphates normalized (normalized SPAP) to transfection levels of a control gene. Results from these experiments are depicted in Figures 1-8. Xenobiotics and steroids were assessed at a concentration of 10 μM unless otherwise indicated on the graph. Bile acids were assessed at a concentration of 100 μM unless otherwise indicated on the graph.

As can be seen from these experiments, PXRs from different species exhibited varying activation patterns in the presence of the same compounds. Similar activation screening assays to these can be performed with other test compounds, most preferably new drugs in development. Results from these

assays are useful in PXR comparative pharmacology and selecting appropriate animal models for preclinical studies predictive of effects in humans.

The following nonlimiting examples further illustrate the present invention.

EXAMPLES

Example 1: Characterization of the sequence

Four PXR LBD sequences from pig, dog, zebrafish, and rhesus monkey were cloned. The isolation of each sequence was achieved using essentially the same strategy for each. A small stretch of the LBD was obtained using either cross-hybridizing PCR primers from another species, or by finding some portion of the LBD sequence in the EST database. The remainder of the LBD was subsequently isolated by PCR amplification of flanking sequence using a primer from within the starting sequence combined with either A) a degenerate oligo representing the canonical P-box of the DBD to isolate 5' sequence, or B) oligo $d(T)_{20}$-G, $d(T)_{20}$-C, or $d(T)_{20}$-A to isolate 3' sequence. After deriving the sequence to the poly (A) tail, the full-length LBD was produced using primers flanking the coding sequence. Wild-type sequence was determined through examination of at least three independent amplifications of each full-length LBD.

To clone pig PXR LBD, total mRNA was prepared from frozen pig liver (1g) using the FastTrack 2.0 RNA Preparation kit (InVitrogen, San Diego, CA). Oligo d(T)-primed cDNA synthesis was carried out by RT-PCR using a cDNA Cycle Kit (InVitrogen, San Diego, CA). An approximately 250 base pair stretch of pig PXR LBD was amplified from this cDNA using homologous mouse PXR LBD primers.

To clone dog PXR LBD, human PXR LBD primers were used to amplify an approximately 450 base pair fragment from a dog liver 5'-stretch λgt11 cDNA library (Clontech, Palo Alto, CA).

To clone rhesus PXR LBD, human primers were used to amplify all but the termini of the rhesus PXR LBD from a rhesus liver cDNA library. Primer sequences 5'-TGC CGT GTA TGT GGG GAC AAG GC-3' (SEQ ID NO:9) and 5'-GGC ATG AAG AA GAG ATG ATC ATG-3' (SEQ ID NO:10) were used to amplify a 274 base pair fragment from a Rhesus liver library constructed in the CMVSport6 vector. The fragment was sequenced and new primers were designed based on this Rhesus sequence. These primers were then used with vector arm primers to amplify the regions 5' and 3' of the known sequenced fragment.

To clone zebrafish (*Danio rerio*) PXR LBD, an initial fragment of the PXR LBD was identified as an EST sequence (Accession # AI943313). This sequence was used to design primers for amplification of the entire LBD from cDNA synthesized using zebrafish embryo (48h) oligo d(T)-primed cDNA.

Example 2: Characterization of Selected PXRs via Cotransfection Assays

In order to assess ligands for their ability to activate PXR in a cell-based assay, a transient transfection approach was utilized. PXR ligand binding domains were fused to the Gal4 DNA binding domain and tested against a reporter gene regulated by a Gal4 response element (from the yeast UAS_G). Lipofectamine-based transient transfection assays were conducted as described previously (Jones et al., 2000 Molecular Endocrinology, volume 14, pp. 27-39), except that a UAS-tk-SPAP reporter vector was used instead of the (CYP3A1 $DR3)_2$-tk-CAT reporter vector. When testing bile acids, an

expression plasmid encoding intestinal bile acid transporter (IBAT) was added to facilitate cellular uptake of bile acids.

What is claimed is:

1. An isolated pregnane X nuclear receptor polypeptide comprising:

(a) an amino acid sequence of SEQ ID NO: 2, 4, 6 or 8;

(b) a variant of the amino acid sequence as defined in (a) which modulates P450 3A4 levels or activity; or

(c) a fragment of (a) or (b) which modulates P450 3A4 levels or activity.

2. A polypeptide according to claim 1 wherein the variant (b) has at least 80% identity to the amino acid sequence of SEQ ID NO: 2, 4, 6 or 8.

3. A polynucleotide encoding a polypeptide according to claim 1.

4. A polynucleotide according to claim 3 which is a cDNA sequence.

5. A polynucleotide encoding a pregnane X receptor polypeptide which modulates P450 3A4 levels or activity, said polynucleotide comprising:

(a) a nucleic acid sequence of SEQ ID NO: 1, 3, 5 or 7;

(b) a nucleic acid sequence which hybridizes under stringent conditions to the nucleic acid sequence as defined in (a);

(c) a nucleic acid sequence that is degenerate as a

result of the genetic code to the nucleic acid sequence as defined in (a) or (b); or

(d) a nucleic acid sequence having at least 60% identity to the nucleic acid sequence as defined in (a), (b) or (c).

6. The polynucleotide of claim 3 wherein the polynucleotide encodes amino acids 106 to 434 set forth in SEQ ID NO: 4.

7. The polynucleotide of claim 3 wherein the polynucleotide encodes amino acids 41 to 105 set forth in SEQ ID NO:4.

8. The polynucleotide of claim 3 wherein the polynucleotide encodes amino acids 1 to 40 set forth in SEQ ID NO:4.

9. A fusion protein comprising:

(a) a DNA binding or ligand binding domain of the pregnane X receptor of claim 1; and

(b) a non-pregnane X receptor-derived amino acid sequence.

10. An isolated polynucleotide encoding the fusion protein of claim 9.

11. An expression vector comprising a polynucleotide according to any one of claims 3 to 8 or 10.

12. A host cell comprising an expression vector according to claim 11.

13. An antibody specific for a polypeptide according to claim 1.

14. A method for the identification of a compound that modulates pregnane X receptor activity and/or expression, said method comprising:

(a) contacting a test compound with a canine, porcine, primate or Zebrafish pregnane X recpetor polypeptide or polynucleotide; and

(b) determining an effect of the test compound on the activity and/or expression of said polypeptide or polynucleotide.

15. A method according to claim 14 wherein the polypeptide is expressed in a cell.

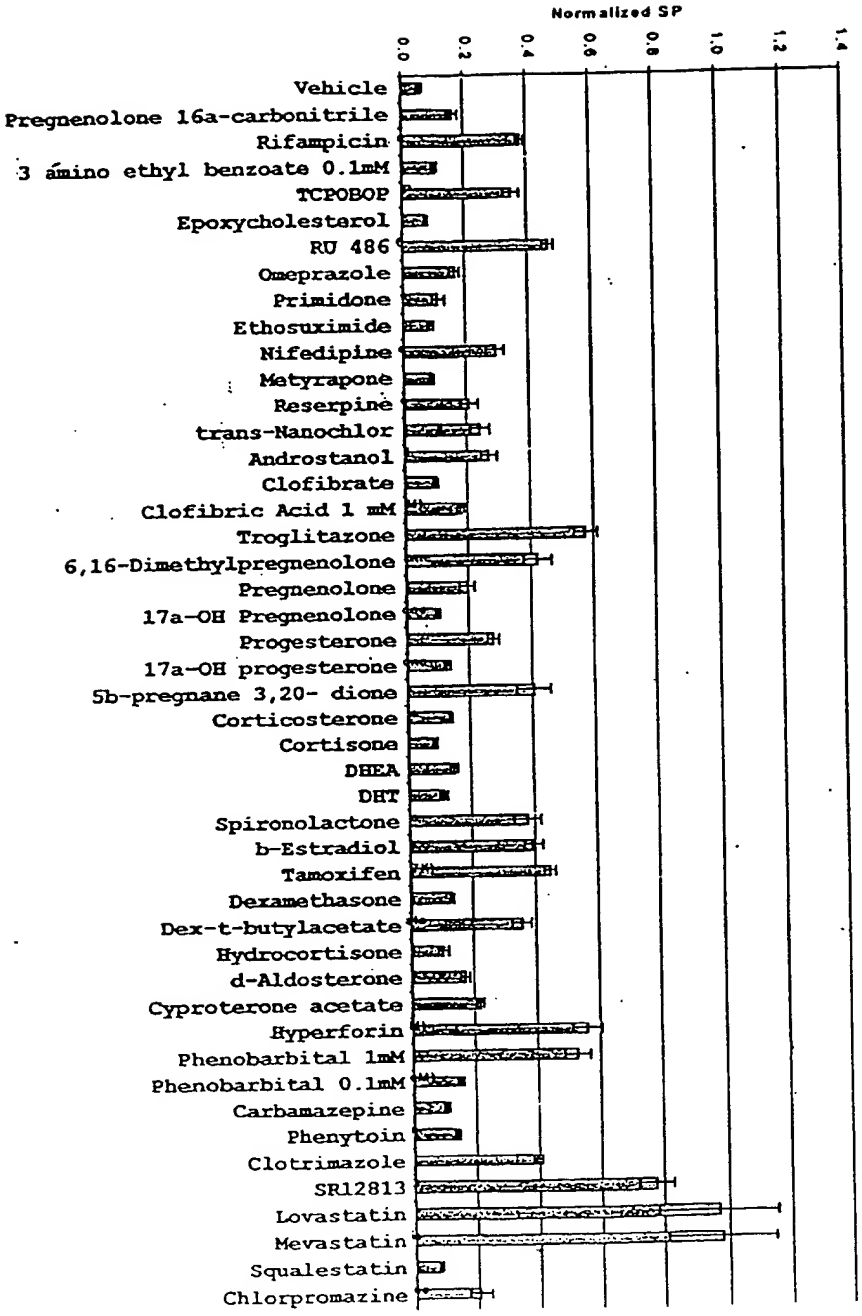
16. A substance which modulates pregnane X receptor activity and which is identifiable by a method according to claim 14 or 15.

17. A non-human transgenic animal expressing a PXR polypeptide of claim 1 or a mutant thereof.

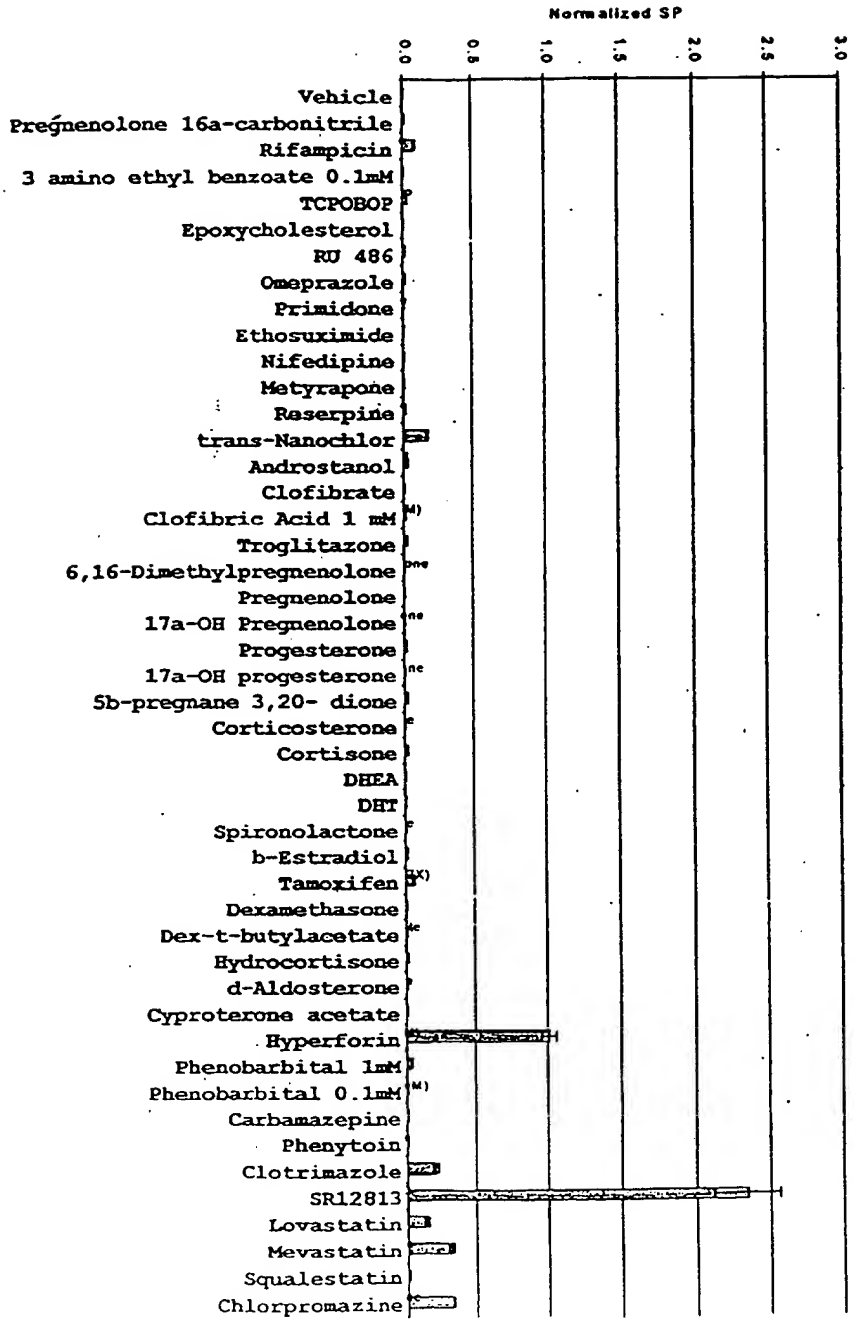
18. A method for selecting a preclinical animal model which is predictive of affects of a test compound on P450 3A4

in humans comprising comparing *in vitro* activation of human PXR in the presence of the test compound with *in vitro* activation of PXRs from preclinical animal models in the presence of the test compound, wherein a PXR from a preclinical animal model exhibiting similar *in vitro* activation to the human PXR is indicative of the preclinical animal model being predictive of the affects of the test compound on P450 3A4 in humans.

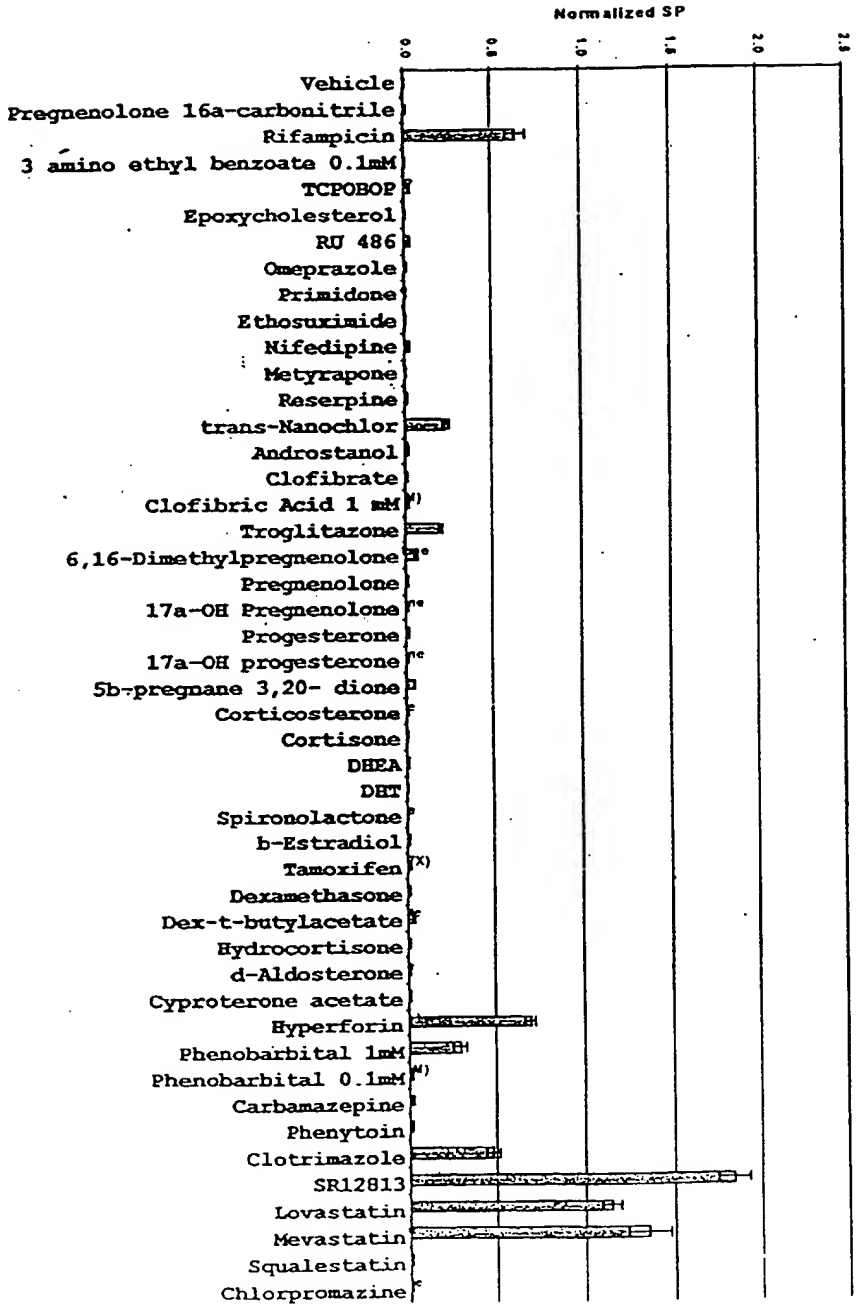
19. The method of claim 18 wherein the pregnane X receptor from the preclinical model is a canine, porcine, primate or Zebrafish pregnane X receptor.



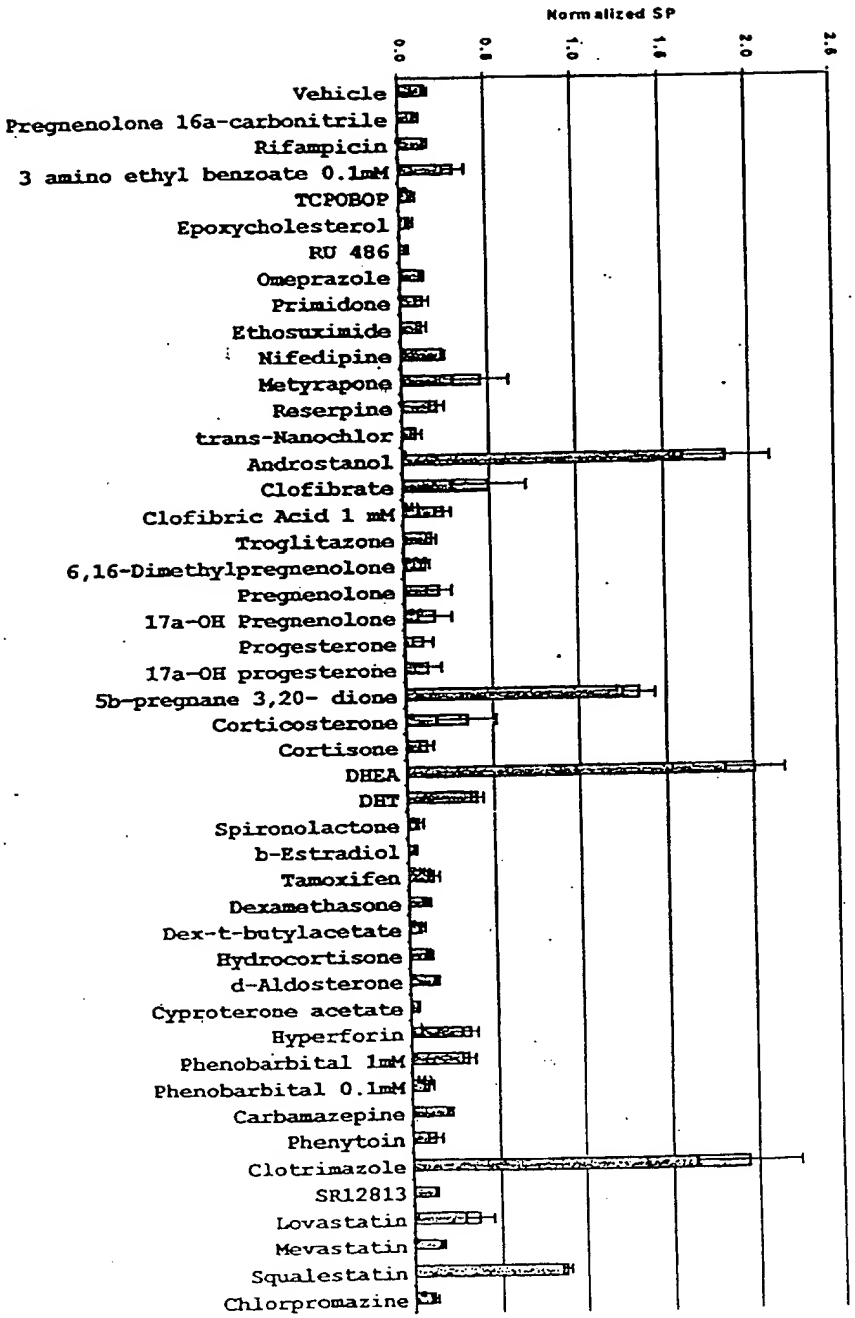
Rhesus GAL4 PXR Activation by Steroids/Xenobiotics
(All compounds are 10 μM unless otherwise noted)



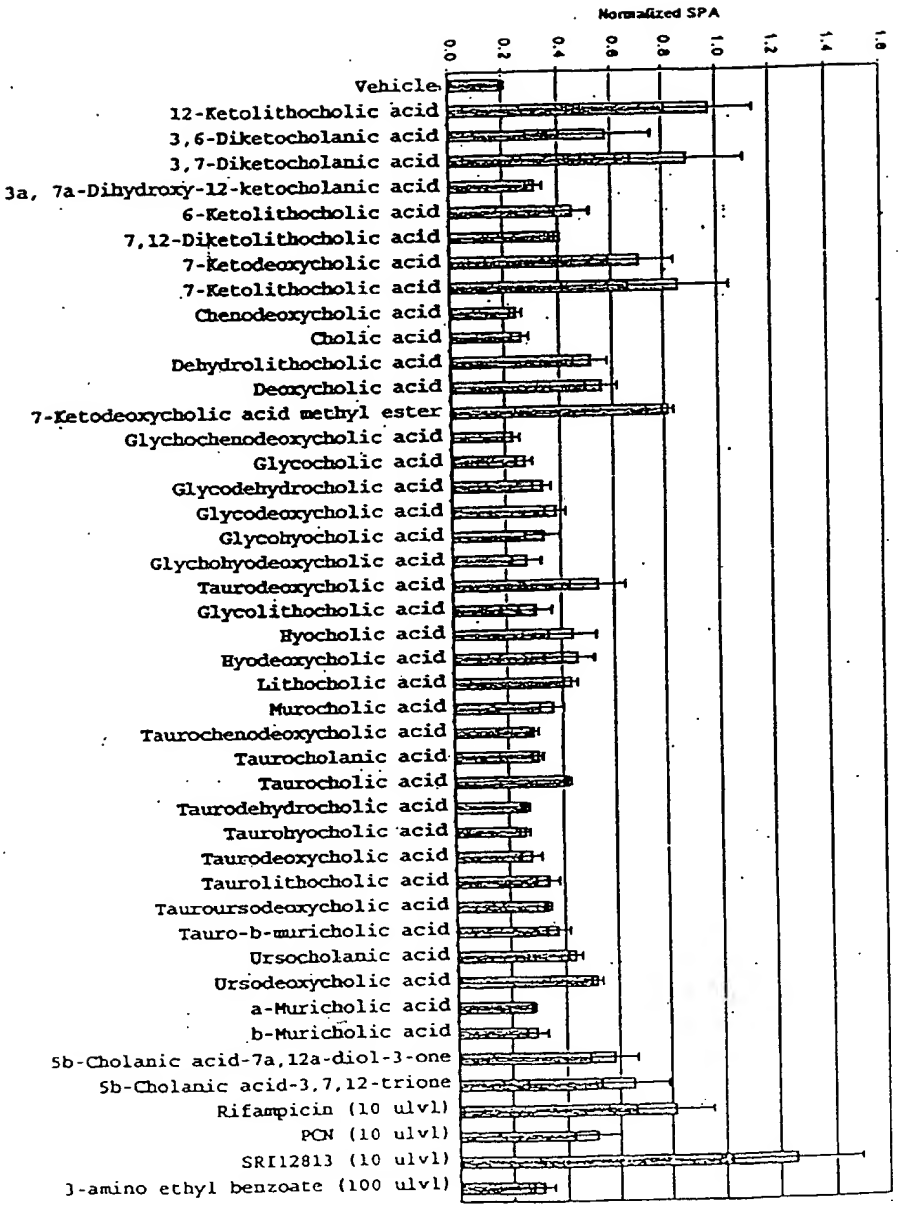
Dog GAL4 PXR Activation by Steroids/Xenobiotics
(All compounds are 10 μM unless otherwise noted)



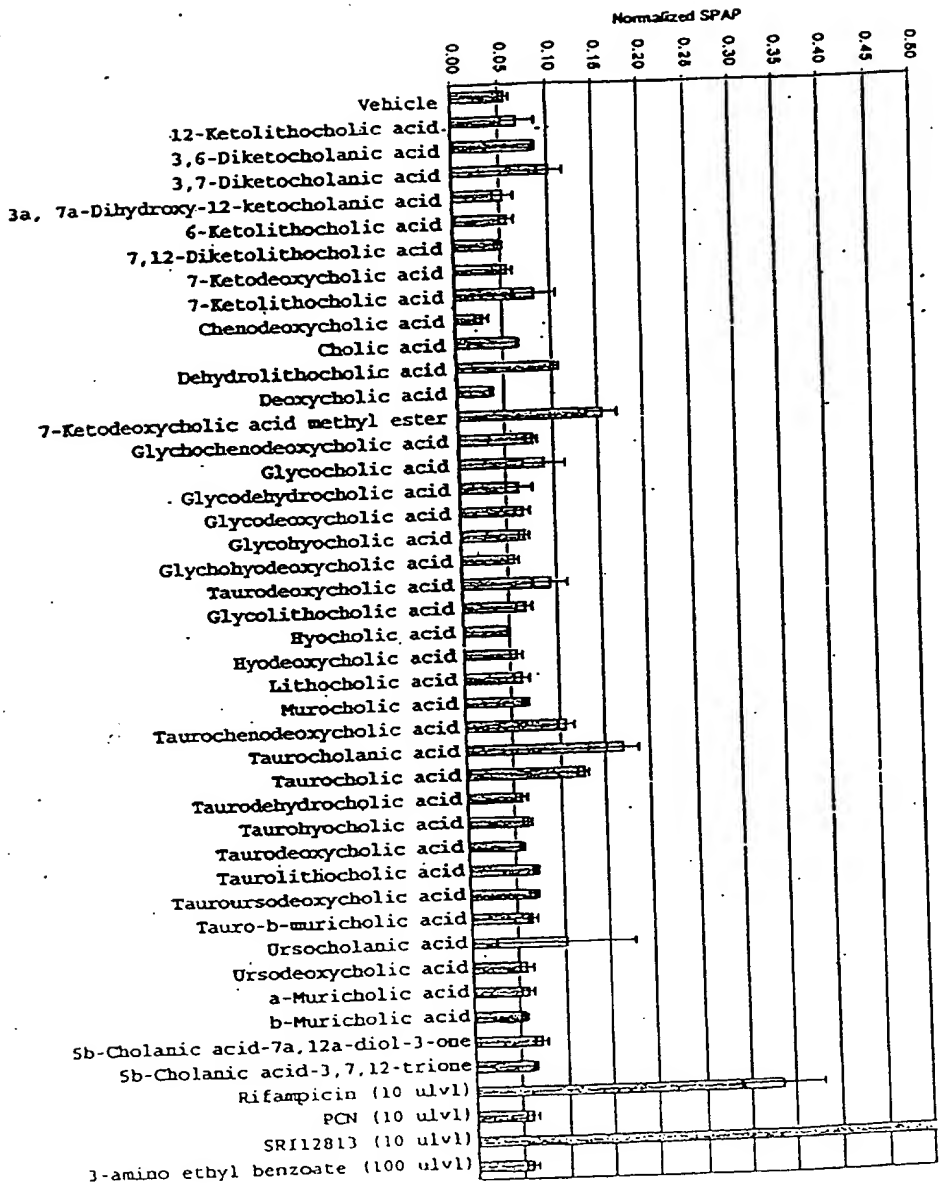
Pig GAL4 PXR Activation by Steroids/Xenobiotics
(All compounds are 10 μM unless otherwise noted)



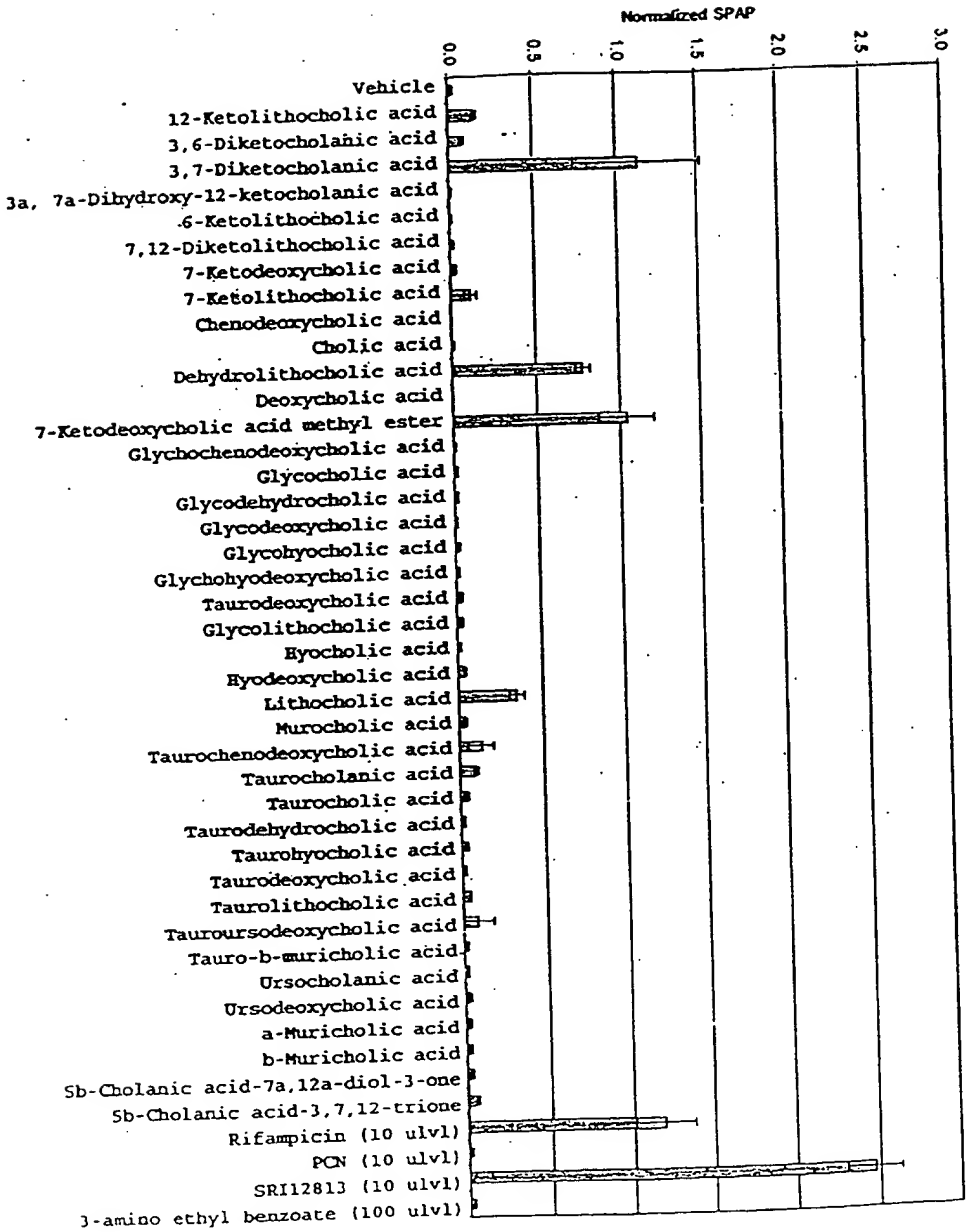
Zebrafish GAL4 PXR Activation by Sterolds/Xenobiotics
(All compounds are 10 μM unless otherwise noted)



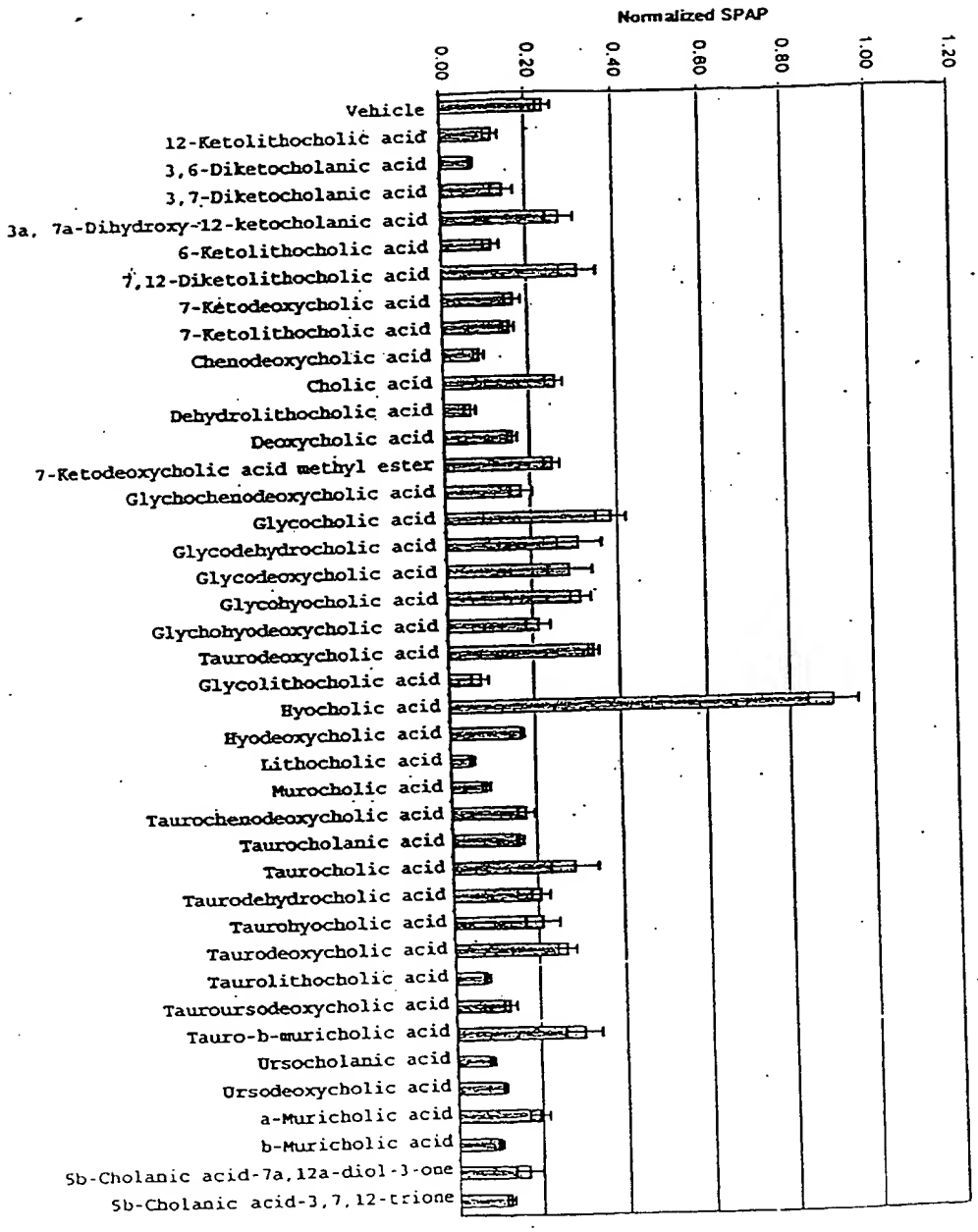
Rhesus GAL4 PXR Activation by Bile Acids
(All compounds are 100 μM unless otherwise noted)
Rifampicin, PCN, SR12813, and 3 amino ethyl benzoate added as controls



Dog GAL4 PXR Activation by Bile Acids
(All compounds are 100 μM unless otherwise noted)
Rifampicin, PCN, SR12813, and 3 amino ethyl benzoate added as controls



Plg GAL4 PXR Activation by Bile Acids
(All compounds are 100 µM unless otherwise noted)
Rifamplcln, PCN, SR12813, and 3 amino ethyl benzoate added as controls



Zebrafish GAL4 PXR Activation by Bile Acids
(All compounds are 100 μM unless otherwise noted)

Nucleotide and Amino Acid Sequence of Canine PXR Ligand Binding Domain (SEQ ID NO:1)

```
GGCATGAAGAAGGAGATGATCATGTCCGACGCGGCTGTGGAGCAGAGGCGGGCTCTGATCCGGAGGAAAAAG
 G  M  K  K  E  M  I  M  S  D  A  A  V  E  Q  R  R  A  L  I  R  R  K  K
CGAGAACGGATGGGCGCGTCGCCGCTGGGAGCCAAGGGGCTGAGTGAGGAGCAGCAGACGATGATCCGAGAG
 R  E  R  M  G  A  S  P  L  G  A  K  G  L  S  E  E  Q  Q  T  M  I  R  E
CTGATGGATGCCCAGATGAAAACCTTTGACACCACCTTCTCCAACTTCAAGGATTTCCGGCTGCCGGCCGCG
 L  M  D  A  Q  M  K  T  F  D  T  T  F  S  N  F  K  D  F  R  L  P  A  A
TGCAGCAGCGGGCGCGAGGTCCCAGGAGCGGCGCACACTCCAGTGGGGGAGGAAGCTGCCAAGTGGAGCCAG
 C  S  S  G  R  E  V  P  G  A  A  H  T  P  V  G  E  E  A  A  K  W  S  Q
GTCAGGGAGGATCTGTGCTCGCTGAAGGTGTGCCTGCGGCTGCGCGGGGAGGACGGCAGCGTCCAGAACTAC
 V  R  E  D  L  C  S  L  K  V  C  L  R  L  R  G  E  D  G  S  V  Q  N  Y

ACACCCCAGGCCGACCGCAGCGGCGCCGAGATCTTTTCCCTGCTGCCCCACATGGCTGACATGTCCACCTAC
 T  P  Q  A  D  R  S  G  A  E  I  F  S  L  L  P  H  M  A  D  M  S  T  Y

ATGTTCAAAGGCGTCATCAACTTTGCCAAAGTCATCTCCCACTTCAGGGAATTGCCCATCGAGGACCAGATC
 M  F  K  G  V  I  N  F  A  K  V  I  S  H  F  R  E  L  P  I  E  D  Q  I
TCGCTGCTAAAGGGGGCCACCTTCGAGGTGTGCCAGCTGAGGTTCAACACGGTGTTCAACGCAGAGACCGGA
 S  L  L  K  G  A  T  F  E  V  C  Q  L  R  F  N  T  V  F  N  A  E  T  G
ACCTGGGAGTGTGGCCGGCTGTCCTACTGCCTGGAAGACCCTGCAGGCGGCTTCCAGCAGCTTCTCCTGGAG
 T  W  E  C  G  R  L  S  Y  C  L  E  D  P  A  G  G  F  Q  Q  L  L  L  E
CCAGTGCTGAAGTTCCACTACAGGCTGAAGAGGCTGCAGCTGCATAAGGAGGAGTATGTGCTGATGCAGGCC
 P  V  L  K  F  H  Y  R  L  K  R  L  Q  L  H  K  E  E  Y  V  L  M  Q  A
ATCTCTCTTTTCTCCCCAGACCGCCCAGGTGTGGTGCAGCGCAGCGTGGTGGACCAGCTGCAGGAGAGATTT
 I  S  L  F  S  P  D  R  P  G  V  V  Q  R  S  V  V  D  Q  L  Q  E  R  F
GCCATCGCCCTGAAGGCCTACATCGAGTGCAATCGGCCGCAGCCTGCCCACCGGTTCCTGTTCCTGAAGATC
 A  I  A  L  K  A  Y  I  E  C  N  R  P  Q  P  A  H  R  F  L  F  L  K  I

ATGGCCATGCTCACCGAGCTTCGCAGCATCAATGCCCAGCACACCCAGAAGCTGCTGCGCATCCAGGACATA
 M  A  M  L  T  E  L  R  S  I  N  A  Q  H  T  Q  K  L  L  R  I  Q  D  I
CACCCCTTCGCCAGCCCCCTCATGCAGGAGCTGTTCAGCATCACGGACGGCTGA
 H  P  F  A  S  P  L  M  Q  E  L  F  S  I  T  D  G  *
```

Amino Acid Sequence of Canine PXR Ligand Binding Domain (SEQ ID NO:2)

```
GMKKEMIMSD AAVEQRRALI RRKKRERMGA SPLGAKGLSE EQQTMIRELM
```

DAQMKTFDTT FSNFKDFRLP AACSSGREVP GAAHTPVGEE AAKWSQVRED
LCSLKVCLRL RGEDGSVQNY TPQADRSGAE IFSLLPHMAD MSTYMFKGVI
NFAKVISHFR ELPIEDQISL LKGATFEVCQ LRFNTVFNAE TGTWECGRLS
YCLEDPAGGF QQLLLEPVLK FHYRLKRLQL HKEEYVLMQA ISLFSPDRPG
VVQRSVVDQL QERFAIALKA YIECNRPQPA HRFLFLKIMA MLTELRSINA
QHTQKLLRIQ DIHPFASPLM QELFSITDG

Nucleotide and Amino Acid Sequence of Full-Length Primate PXR (SEQ ID NO:3)

TCCTTGGTAAAGCTACTCCTTGATCGATCCTTTGCACCTGATTGTTCAAAGTGGACCCCAGGGG

GAAGTCAGAGCGAAGAACTTACCGCTAAGCAGTCCAAGAGGCCCAGAAGCAAAC

CTGGAGGTGAGACCCAAAGAAGGCTGGAACCATGCTGACTTTGTATACTGTGAGGACACAGAGTTTGCTCCT
 L E V R P K E G W N H A D F V Y C E D T E F A P
GGAAAGCCCACTGTCAACGCAGATGAGGAAGTTGGGGGTCCCCAAATCTGCCGTGTATGTGGGGACAAGGCC
 G K P T V N A D E E V G G P Q I C R V C G D K A
ACTGGTTATCACTTCAATGTCATGACATGTGAAGGATGCAAGGGCTTTTTCAGGAGGGCCATGAAACGCAAC
 T G Y H F N V M T C E G C K G F F R R A M K R N
GCCCGCCTTAGGTGCCCCTTCCGGAAGGGCGCCTGCGAGATCACCCGGAAGACCCGGCGACAGTGCCAGGCC
 A R L R C P F R K G A C E I T R K T R R Q C Q A
TGCCGGCTGCGCAAGTGCCTGGAGAGCGGCATGAAGAAGGAGATGATCATGTCCGACGCGGCCGTAGAGGAG
 C R L R K C L E S G M K K E M I M S D A A V E E
AGGCGGGCCTTGATCAAGAGGAAGAAAAGAGAACGGATCGGGACTCAGCCACCCGGAGTGCAGGGGCTGACG
 R R A L I K R K K R E R I G T Q P P G V Q G L T
GAGGAGCAGCGGATGATGATCAGGGAGCTGATGGACGCTCAGATGAAAACCTTTGACACTACCTTCTCCCAT
 E E Q R M M I R E L M D A Q M K T F D T T F S H
TTCAAGAATTTCCGGCTGCCAGGGGTGCTTAGCAGTGGCTGTGAGATGCCAGAGTCTCTGCAGGCCCCATCG
 F K N F R L P G V L S S G C E M P E S L Q A P S
AGGGAAGAAGCTGCCAAGTGGAACCAGGTCAGGAAAGATCTGTGGTCTGTGAAGGTCTCCGTGCAGCTGCGG
 R E E A A K W N Q V R K D L W S V K V S V Q L R
GGGGAGGATGGCAGTGTCTGGAACTACAAACCCCCAGCCGACAATGGCGGGAAAGAGATCTTCTCCTTGCTG
 G E D G S V W N Y K P P A D N G G K E I F S L L
CCCCACATGGCTGACATGTCAACCTACATGTTCAAAGGCATCATCAACTTTGCCAAAGTCATCTCCTACTTC
 P H M A D M S T Y M F K G I I N F A K V I S Y F
AGGGACCTGCCCATCGAGGACCAGATCTCCCTACTGAAGGGGGCCACTTTTGAGCTGTGCCAGCTGAGATTC
 R D L P I E D Q I S L L K G A T F E L C Q L R F
AACACAGTATTCAACGCGGAGACTGGAACTTGGGAGTGTGGCCGGCTGTCCTACTGCTTGGAAGACCCTGCA
 N T V F N A E T G T W E C G R L S Y C L E D P A
GGTGGTTTCCAGCAACTTCTGCTGGAGCCCATGCTGAAATTCCACTACATGCTGAAGAAGCTGCAGCTACAC
 G G F Q Q L L L E P M L K F H Y M L K K L Q L H
GAGGAGGAGTATGTGCTGATGCAGGCCATCTCCCTCTTCTCCCCAGACCGCCCAGGTGTGGTGCAGCACCGC
 E E E Y V L M Q A I S L F S P D R P G V V Q H R
GTGGTGGACCAGCTGCAGGAGCAATACGCTATTACTCTGAAGTCCTACATTGAATGCAATCGGCCCCAGCCT
 V V D Q L Q E Q Y A I T L K S Y I E C N R P Q P
GCTCATAGGTTCCTGTTCCTGAAGATCATGGCTATGCTCACCGAGCTCCGCAGCATCAACGCCCAGCACACC
 A H R F L F L K I M A M L T E L R S I N A Q H T
CAGCGGCTGCTGCGCATCCAGGACATACACCCCTTTGCTACGCCCCTCATGCAGGAGTTGTTCGGCATCACG
 Q R L L R I Q D I H P F A T P L M Q E L F G I T
GGTAGCTGAGTGGCTGCCCTTGGGTGA
 G S *

Amino Acid Sequence of Full-Length Primate PXR (SEQ ID NO:4)

LEVRPKEGWNHADFVYCEDTEFAPGKPTVNADEEVGGPQICRVCGDKA
TGYHFNVMTCEGCKGFFRRAMKRNARLRCPFRKGACEITRKTRRQCQA
CRLRKCLESGMKKEMIMSDAAVEERRALIKRKKRERIGTQPPGVQGLT
EEQRMMIRELMDAQMKTFDTTFSHFKNFRLPGVLSSGCEMPESLQAPS
REEAAKWNQVRKDLWSVKVSVQLRGEDGSVWNYKPPADNGGKEIFSLL
PHMADMSTYMFKGIINFAKVISYFRDLPIEDQISLLKGATFELCQLRF
NTVFNAETGTWECGRLSYCLEDPAGGFQQLLLEPMLKFHYMLKKLQLH
EEEYVLMQAISLFSPDRPGVVQHRVVDQLQEQYAITLKSYIECNRPQP
AHRFLFLKIMAMLTELRSINAQHTQRLLRIQDIHPFATPLMQELFGIT
GS

Nucleotide and Amino Acid Sequence of Porcine PXR Ligand Binding Domain (SEQ ID NO:5)

GGCATGAGGAAGGAAATGATCATGTCAGATGCAGCTGTGGAGCAGAGGCGGGCCTTGATCAGGAGGAAGAAA
G M R K E M I M S D A A V E Q R R A L I R R K K
CGAGAACAGATCGGGGCTCAGCCCCCAGGAGCCAAGGGTCTCACTGAAGAGCAGCGGACAATGATCAGTGAG
R E Q I G A Q P P G A K G L T E E Q R T M I S E
CTGATGAACGCTCAGATGAAAACCTTTGACACCACCTTCACACATTTCAAGAATTTTCGGTTACCAGAGGTG
L M N A Q M K T F D T T F T H F K N F R L P E V
CTTAGCAGTAGCCTCGAGATTCCAGAGTGTCTGCAGACTCCGTCGTCAAGGGAAGAAGCTGCCAAGTGGAGC
L S S S L E I P E C L Q T P S S R E E A A K W S

AAGCTCAGGGAAGATCTGTGCTCAGTGAAACTCTCTCTGCAGCTAAGGGGGGAAGATGGTAGCGTCTGGAAC
K L R E D L C S V K L S L Q L R G E D G S V W N
TACAAACCCCCAGCAGACAACAGTGGGAAAGAGATCTTTTCCCTGCTGCCCCACATAGCTGACATGTCAACC
Y K P P A D N S G K E I F S L L P H I A D M S T
TACATGTTCAAAGGCATTATCAACTTTGCCAAAGTCATCTCCTACTTCAGGGACTTGCCCATTGAGGACCAG
Y M F K G I I N F A K V I S Y F R D L P I E D Q
ATCTCTCTGCTGAAGGGGGCCACCTTTGAGCTGTGCCAGCTGAGATTCAACACGGTGTTCAACGCAGAGACG
I S L L K G A T F E L C Q L R F N T V F N A E T
GGGACCTGGGAGTGTGGTCGGCTGTCCTACAGCTTGGAAGACCCCTCAGGTGGCTTCCAGCAGCTTCTCCTG
G T W E C G R L S Y S L E D P S G G F Q Q L L L
CAGCCCATGCTGAAATTCCACTACATGCTGAAGAAGCTGCAGCTGCATAAGGAGGAGTATGTGCTGATGCAG
Q P M L K F H Y M L K K L Q L H K E E Y V L M Q
GCCATCTCCCTTTTCTCTCCAGACCGCCCGGGTGTGGTGCAACGCCAAGTGGTGGACCAGCTGCAGGAGAGG
A I S L F S P D R P G V V Q R Q V V D Q L Q E R
TTTGCCATTACCCTGAAGGCCTACATCGAGTGCAACCGGCCCCAGCCTGCCCACCGATTCCTGTTCCTGAAG

F A I T L K A Y I E C N R P Q P A H R F L F L K
ATCATGGCTATGCTCACTGAGCTCCGCAGCATCAACGCCCAACACACCCAGCGGCTGCTGCGAATCCAGGAC
I M A M L T E L R S I N A Q H T Q R L L R I Q D
ATACACCCCTTCGCCACCCCACTCATGCAGGAGTTATTCAGCATCACAGAAAGCTGA
I H P F A T P L M Q E L F S I T E S *

Amino Acid Sequence of Porcine PXR Ligand Binding Domain (SEQ ID NO:6)

GGCATGAGGAAGGAAATGATCATGTCAGATGCAGCTGTGGAGCAGAGGCGGGCCTTGATCAGGAGGAAGAAA
G M R K E M I M S D A A V E Q R R A L I R R K K
CGAGAACAGATCGGGGCTCAGCCCCCAGGAGCCAAGGGTCTCACTGAAGAGCAGCGGACAATGATCAGTGAG
R E Q I G A Q P P G A K G L T E E Q R T M I S E
CTGATGAACGCTCAGATGAAAACCTTTGACACCACCTTCACACATTTCAAGAATTTTCGGTTACCAGAGGTG
L M N A Q M K T F D T T F T H F K N F R L P E V
CTTAGCAGTAGCCTCGAGATTCCAGAGTGTCTGCAGACTCCGTCGTCAAGGGAAGAAGCTGCCAAGTGGAGC
L S S S L E I P E C L Q T P S S R E E A A K W S

AAGCTCAGGGAAGATCTGTGCTCAGTGAAACTCTCTCTGCAGCTAAGGGGGGAAGATGGTAGCGTCTGGAAC
K L R E D L C S V K L S L Q L R G E D G S V W N
TACAAACCCCCAGCAGACAACAGTGGGAAAGAGATCTTTTCCCTGCTGCCCCACATAGCTGACATGTCAACC
Y K P P A D N S G K E I F S L L P H I A D M S T
TACATGTTCAAAGGCATTATCAACTTTGCCAAAGTCATCTCCTACTTCAGGGACTTGCCCATTGAGGACCAG
Y M F K G I I N F A K V I S Y F R D L P I E D Q
ATCTCTCTGCTGAAGGGGGCCACCTTTGAGCTGTGCCAGCTGAGATTCAACACGGTGTTCAACGCAGAGACG
I S L L K G A T F E L C Q L R F N T V F N A E T
GGGACCTGGGAGTGTGGTCGGCTGTCCTACAGCTTGGAAGACCCCTCAGGTGGCTTCCAGCAGCTTCTCCTG
G T W E C G R L S Y S L E D P S G G F Q Q L L L
CAGCCCATGCTGAAATTCCACTACATGCTGAAGAAGCTGCAGCTGCATAAGGAGGAGTATGTGCTGATGCAG
Q P M L K F H Y M L K K L Q L H K E E Y V L M Q
GCCATCTCCCTTTTCTCTCCAGACCGCCCGGGTGTGGTGCAACGCCAAGTGGTGGACCAGCTGCAGGAGAGG
A I S L F S P D R P G V V Q R Q V V D Q L Q E R
TTTGCCATTACCCTGAAGGCCTACATCGAGTGCAACCGGCCCCAGCCTGCCCACCGATTCCTGTTCCTGAAG
F A I T L K A Y I E C N R P Q P A H R F L F L K
ATCATGGCTATGCTCACTGAGCTCCGCAGCATCAACGCCCAACACACCCAGCGGCTGCTGCGAATCCAGGAC
I M A M L T E L R S I N A Q H T Q R L L R I Q D
ATACACCCCTTCGCCACCCCACTCATGCAGGAGTTATTCAGCATCACAGAAAGCTGA
I H P F A T P L M Q E L F S I T E S *
*

Nucleotide and Amino Acid Sequence of Zebrafish PXR Ligand Binding Domain (SEQ ID NO:7)

GGCATGAAGAGAGAGCTGATCATGTCGGATGAGGCGGTGGAGAAGCGGAGGTTGCAGATCAGGAGGAAGAGG
G M K R E L I M S D E A V E K R R L Q I R R K R
ATGCAGGAAGAGCCTGTAACTCTCACTCCTCAACAGGAAGCTGTCATACAAGAGCTGCTYAACGCACACAAG
M Q E E P V T L T P Q Q E A V I Q E L L N A H K
AAAACCTTCGACATGACTTGTGCCCATTTCAGTCAGTTCCGGCCTTTAGATCGGGATCAGAAGTCTGTGTCC
K T F D M T C A H F S Q F R P L D R D Q K S V S
GAGTCGAGTCCACTCACAAACGGCAGCTGGATCGATCACAGACCCATCGCTGAAGACCCAATGCAGTGGGTC
E S S P L T N G S W I D H R P I A E D P M Q W V
TTCAATCCCACTTCGCTCTCGTCCTCTTCCTCCAGCTACCAGAGCCTTGACAATAAAGAGAAGAAGCACTTT
F N P T S L S S S S S S Y Q S L D N K E K K H F
AAAAGTGGCAACTTCTCCTCTCTGCCACACTTCACAGACCTCACCACGTACATGATCAAGAATGTCATCAAC
K S G N F S S L P H F T D L T T Y M I K N V I N
TTCGGGAAGACGCTGACAATGTTTAGGGCTCTGGTTATGGAGGACCAGATCTCGCTGCTGAAAGGTGCCACC
F G K T L T M F R A L V M E D Q I S L L K G A T
TTTGAGATCATTCTGATTCACTTCAACATGTTCTTTAATGAAGTGACGGGAATTTGGGAGTGCGGCCCCCTG
F E I I L I H F N M F F N E V T G I W E C G P L
CAGTACTGCATGGATGATGCCTTTCGAGCTGGTTTTCAGCACCATCTGCTGGACCCAATGATGAATTTCCAT
Q Y C M D D A F R A G F Q H H L L D P M M N F H
TACACACTGCGTAAGCTGCGTTTGCATGAGGAGGAGTATGTGCTGATGCAGGCCCTCTCTCTCTTTTCACCA
Y T L R K L R L H E E E Y V L M Q A L S L F S P
GATCGCCCTGGTGTGACAGACCACAAAGTGATCGACCGCAACCAGGAAACACTAGCGCTTACCCTAAAGACT
D R P G V T D H K V I D R N Q E T L A L T L K T
TACATTGAGGCCAAGAGAAATGGGCCAGAAAAACATCTGCTGTTCCCAAAGATTATGGGGTGCCTGACCGAG
Y I E A K R N G P E K H L L F P K I M G C L T E
ATGAGGAGCATGAACGAAGAGTACACCAAACAAGTGCTGAAAATCCAGGACATGCAGCCTGAAGTGTCTCCA
M R S M N E E Y T K Q V L K I Q D M Q P E V S P
CTTTGGTTGGAAATAATAAGCAAAGACACCTAAGTCCTCACAAAGCACGTGCAACTCTACTTATTCCCAACT
L W L E I I S K D T *

Amino Acid Sequence of Zebrafish PXR Ligand Binding Domain (SEQ ID NO:8)

GMKRELIMSDEAVEKRRLQIRRKRMQEEPVTLTPQQEAVIQELLNAHK
KTFDMTCAHFSQFRPLDRDQKSVSESSPLTNGSWIDHRPIAEDPMQWV

FNPTSLSSSSSSYQSLDNKEKKHFKSGNFSSLPHFTDLTTYMIKNVIN
FGKTLTMFRALVMEDQISLLKGATFEIILIHFNMFFNEVTGIWECGPL
QYCMDDAFRAGFQHHLLDPMMNFHYTLRKLRLHEEEYVLMQALSLFSP
DRPGVTDHKVIDRNQETLALTLKTYIEAKRNGPEKHLLFPKIMGCLTE
MRSMNEEYTKQVLKIQDMQPEVSPLWLEIISKDT*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/16445

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C07K 14/00; C12N 5/10, 15/11, 15/12, 15/62, 15/63
US CL : 435/69.1, 320.1, 325; 530/350; 536/23.5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/69.1, 320.1, 325; 530/350; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 99/35246 A1 (THE SALK INSTITUTE FOR BIOLOGICAL STUDIES) 15 July 1999 (15.07.99) see entire document.	1-12
A	WO 99/48915 A1 (GLAXO GROUP LIMITED) 30 September 1999 (30.09.99) see entire document.	1-12

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 September 2002 (25.09.2002)	18 OCT 2002
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Michael Pak Telephone No. 703-308-1235

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/16445

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-12

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-12, drawn to a polypeptide, polynucleotide, fusion protein, an expression vector, and a host cell.

Group II, claim(s) 13, drawn to an antibody.

Group III, claim(s) 14-15, drawn to a method for identification of compound.

Group IV, claim(s)16, drawn to a substance which modulates.

Group V, claim(s) 17, drawn to a non-human transgenic animal.

Group VI, claim(s) 18-19, drawn to a method for seleting a preclinical aminoa model.

The International Application does not comply with the requirements of unity of invention (Rules 13.1, 13.2 and 13.3) for the reasons indicated below:

The inventions listed as Groups I-XII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is a polypeptide, polynucleotide, fusion protein, an expression vector, and a host cell. Pursuant to 37 CFR 1.475(d), these claims are considered by the ISA/US to constitute the main invention, and none of the related groups II-VI correspond to the main invention.

The products of Group II, IV and V does not share the same or corresponding special technical feature with Group I, because they are drawn to products having materially different structures and functions, and each defines a separate invention over the art.

The methods of Groups III and VI, do not share the same or corresponding special technical feature with Group I, because the methods have materially different process steps and are practiced for materially different purposes, and each defines a separate invention over the art.

Since Groups I-VI do not share a special technical feature, unity of invention is lacking.

Continuation of B. FIELDS SEARCHED Item 3:

INTERNATIONAL SEARCH REPORT

PCT/US02/16445

BRS, GENESEQ, PIR, SWISSPROT, SPTREMBL

search terms: Pregnane X receptor, cytochrome P450, xenobiotics, P450 3A4 monooxygenase, CYP3A4, RIFAMPICIN

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 November 2002 (28.11.2002)

PCT

(10) International Publication Number
WO 02/095652 A1

(51) International Patent Classification[7]: G06F 19/00

(21) International Application Number: PCT/US02/15701

(22) International Filing Date: 16 May 2002 (16.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/292,053 : 18 May 2001 (18.05.2001) US

(71) Applicants *(for all designated States except US)*: **SMITHKLINE BEECHAM CORPORATION** [US/US]; One Franklin Plaza, Philadelphia, PA 19101 (US). **UNIVERSITY OF NORTH CAROLINA** [US/US]; Office of Technology Development, University of North Carolina at Chapel Hill, 308 Bynum Hall, Campus Box 4105, Chapel Hill, NC 27599 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **KLIEWER, Steven** [US/US]; 3453 Potomac Avenue, Dallas, TX 75205 (US). **REDINBO, Matthew, R** [US/US]; c/o University of North Carolina at Chapel Hill, 308 Bynum Hall, Chapel Hill, NC 27599 (US). **WATKINS, Ryan, E.** [US/US]; c/o University of North Carolina at Chapel Hill, 308 Bynum Hall, Chapel Hill, NC 27599 (US). **WISELY, George, Bruce** [US/US]; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US). **WILLIAMS, Shawn, P.** [US/US]; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 (US).

(74) Agents: **LEVY, David, J.**; GlaxoSmithKline, Five Moore Drive, PO Box 13398, Research Triangle Park, NC 27709 et al. (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/095652 A1

(54) Title: CRYSTALLIZED HUMAN XENOBIOTIC NUCLEAR RECEPTOR PXR/SXR LIGAND BINDING DOMAIN POLYPEPTIDE AND SCREENING METHODS EMPLOYING SAME

(57) Abstract: A solved three-dimensional crystal structure of a human PXR ligand binding domain polypeptide is disclosed, along with a crystal form of the PXR ligand binding domain. Orientations of the ligand SR12813 in the binding cavity are also disclosed. Additionally, methods of designing modulators of the biological activity of PXR, and other PXR ligand binding domain polypeptides, are also disclosed.

Description

CRYSTALLIZED HUMAN XENOBIOTIC NUCLEAR RECEPTOR PXR/SXR LIGAND BINDING DOMAIN POLYPEPTIDE AND SCREENING METHODS EMPLOYING SAME

Technical Field

The present invention relates generally to the structure of the ligand binding domain of PXR, and more particularly to the crystalline structure of the ligand binding domain of PXR. The invention further relates to methods by which modulators and ligands of PXR can be identified.

Abbreviations

ATP	adenosine triphosphate
ADP	adenosine diphosphate
APS	Advanced Photon Source
BSA	bovine serum albumin
CBP	CREB-binding protein
cDNA	complementary DNA
CI	chemical ionization
CPS	counts per second
CYP	cytochrome P450
DBD	DNA binding domain
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EI	electron impact ionization
ER	estrogen receptor
FRET	fluorescent resonance energy transfer
GC	gas chromatography
GC/MS	gas chromatography/mass spectrometry
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic

	acid
hPXR	human pregnane X receptor
HRE	hormone response element
kDa	kilodalton(s)
LBD	ligand binding domain
MAD	multiwavelength anomalous diffraction
MODY	mature onset diabetes of the young
MS	mass spectrometry
NDP	nucleotide diphosphate
nt	nucleotide
NTP	nucleotide triphosphate
PAGE	polyacrylamide gel electrophoresis
PAR	pregnane-activated receptor
PCN	pregnenolone 16a-carbonitrile
PCR	polymerase chain reaction
pI	isoelectric point
PPAR	peroxisome proliferator-activated receptor
PR	progesterone receptor
PXR	pregnane X receptor
RAR	retinoic acid receptor
RIF	rifampicin
RMSD	root-mean-square deviation
RXR	retinoid X receptor
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIRAS	single isomorphous replacement anomalous scattering
SR	SR12813
SXR	steroid and xenobiotic receptor
TIC	total ion chromatogram
TR	thyroid hormone receptor

TTR	plasma transthyretin
VDR	vitamin D receptor
vHNF	variant hepatocyte nuclear factor
WT	wildtype

Amino Acid Abbreviations

Single-Letter Code	Three-Letter Code	Name
A	Ala	Alanine
V	Val	Valine
L	Leu	Leucine
I	Ile	Isoleucine
P	Pro	Proline
F	Phe	Phenylalanine
W	Trp	Tryptophan
M	Met	Methionine
G	Gly	Glycine
S	Ser	Serine
T	Thr	Threonine
C	Cys	Cysteine
Y	Tyr	Tyrosine
N	Asn	Asparagine
Q	Gln	Glutamine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
K	Lys	Lysine
R	Arg	Arginine
H	His	Histidine

Functionally Equivalent Codons

Amino Acid			Codons
Alanine	Ala	A	GCA GCC GCG GCU
Cysteine	Cys	C	UGC UGU

Aspartic Acid	Asp	D	GAC GAU
Glumatic acid	Glu	E	GAA GAG
Phenylalanine	Phe	F	UUC UUU
Glycine	Gly	G	GGA GGC GGG GGU
Histidine	His	H	CAC CAU
Isoleucine	Ile	I	AUA AUC AUU
Lysine	Lys	K	AAA AAG
Methionine	Met	M	AUG
Asparagine	Asn	N	AAC AAU
Proline	Pro	P	CCA CCC CCG CCU
Glutamine	Gln	Q	CAA CAG
Threonine	Thr	T	ACA ACC ACG ACU
Valine	Val	V	GUA GUC GUG GUU
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAC UAU
Leucine	Leu	L	UUA UUG CUA CUC CUG CUU
Arginine	Arg	R	AGA AGG CGA CGC CGG CGU
Serine	Ser	S	ACG AGU UCA UCC UCG UCU

Background Art

Nuclear receptors represent a superfamily of proteins that specifically bind a physiologically relevant small molecule, such as a hormone or vitamin. As a result of a molecule binding to a nuclear receptor, the nuclear receptor changes the ability of a cell to transcribe DNA, i.e. nuclear receptors modulate the transcription of DNA. However they can also have transcription independent actions.

Unlike integral membrane receptors and membrane-associated receptors, nuclear receptors reside in either the cytoplasm or nucleus of eukaryotic cells. Thus nuclear receptors comprise a class of intracellular,

soluble ligand-regulated transcription factors. Nuclear receptors include but are not limited to receptors for glucocorticoids, androgens, mineralcorticoids, progestins, estrogens, thyroid hormones, vitamin D retinoids, and icosanoids. Many nuclear receptors, identified by either sequence homology to known receptors (See, Drewes et al., (1996) *Mol. Cell. Biol.* 16:925-31) or based on their affinity for specific DNA binding sites in gene promoters (See, Sladek et al., *Genes Dev.* 4:2353-65), have unascertained ligands and are therefore termed "orphan receptors".

Structurally, nuclear receptors are generally characterized by two distinct structural elements. First, nuclear receptors comprise a central DNA binding domain that targets the receptor to specific DNA sequences, which are known as hormone response elements (HREs). The DNA binding domains of these receptors are related in structure and sequence, and are located within the middle of the receptor. Second, the C-terminal region of nuclear receptors encompasses the ligand binding domain (LBD). Upon binding a ligand, the receptor shifts to a transcriptionally active state.

The cytochrome P450 (CYP) family of heme-containing proteins play critical roles in the oxidative metabolism of drugs and other xenobiotics in the liver and small intestine. In particular, the CYP3A gene products bind and hydroxylate a wide variety of chemical structures, including >50% of all drugs. Maurel, (1996) in Cytochrome P450: Metabolic and Toxicological Aspects (Ionnides, ed.) CRC Press, Inc., Boca Raton, Florida, 241-70. Expression of CYP3A is induced at the level of transcription by a variety of xenobiotics, including many that are metabolized by CYP3A.

The pregnane X receptor (PXR; NR112), a member of the nuclear receptor family of ligand-activated transcription factors, is a key regulator of CYP3A gene expression in mammalian liver and small intestine. Kliewer et al., (1998) *Cell* 92: 73-82; Lehmann et al., (1998) *J. Clin. Invest.* 102: 1016-23; Bertilsson et al., (1998) *Proc. Nat. Acad. Sci. U.S.A.* 95: 12208-13; Blumberg et al., (1998) *Gene Dev.* 12:3195-205; Xie et al., (2000) *Nature* 406: 435-39. The human ortholog of PXR is alternatively known as pregnane activated receptor (PAR) and steroid and xenobiotic receptor (SXR).

Bertilsson et al., (1998) *Proc. Nat. Acad. Sci. U.S.A.* 95: 12208-13; Blumberg et al., (1998) *Gene Dev.* 12: 3195-205. PXR is activated by most of the xenobiotics that are known to induce CYP3A gene expression, including commonly used drugs such as the antibiotic rifampicin, the glucocorticoid dexamethasone, and the herbal therapy St. John's wort. Kliewer et al., (1998) *Cell* 92: 73-82; Lehmann et al., (1998) *J. Clin. Invest.* 102: 1016-23; Bertilsson et al., (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95: 12208-13; Blumberg et al., (1998) *Gene Dev.* 12: 3195-205; Jones et al., (2000) *Mol. Endocrinol.* 14: 27-39; Moore et al., (2000) *Proc. Nat. Acad.Sci. U.S.A.* 97: 7500-502.

Like other nuclear receptors, PXR contains both a DNA binding domain and a ligand binding domain (LBD). PXR binds to the xenobiotic response elements in the regulatory regions of CYP3A genes as a heterodimer with the 9-cis retinoic acid receptor (RXR). Kliewer et al., (1998) *Cell* 92: 730-82; Lehmann et al., (1998) *J. Clin. Invest.* 102: 1016-23; Blumberg et al., (1998) *Gene Dev.* 12: 3195-205.

PXR mediates potentially dangerous drug-drug interactions by upregulating CYP3A expression in response to one compound, which in turn can then lead to the metabolism of other drugs vital to survival. For example, patients taking the herbal anti-depressant St. John's wort have exhibited a dramatic drop in serum levels of other critical drugs, including the antiretroviral drug Indinavir (also known as CRIXIVAN®) and the immunosuppressant compound cyclosporin. Piscitelli et al., (2000) *Lancet* 355: 547-48; Ruschitzka et al., (2000) *Lancet* 355: 548-49. The molecular basis of this effect was traced to PXR. Moore et al., (2000) *Proc. Nat. Acad.Sci. U.S.A.* 97: 7500-502; Wentworth et al., (2000) *J. Endocrinol.* 166: R11-R16. Hyperforin, a constituent of St. John's wort, activates PXR and upregulates CYP3A expression, which leads to the metabolism of these vital drugs. Unraveling the molecular basis of PXR activation by ligands could prove critical in preventing such potentially deadly cross-reactivities between drugs.

Unlike the steroid, retinoid, and thyroid hormone receptors, which are highly selective for their cognate hormone, PXR has evolved to detect structurally-diverse compounds. These include drugs and xenobiotics, as well

as endogenous compounds, including the toxic bile acid lithocholic acid and certain C21 steroids (pregnanes). Kliewer et al., (1998) *Cell* 92: 730-82; Bertilsson et al., (1998) *Proc. Nat. Acad. Sci. U.S.A.* 95: 12208-13; Jones et al., (2000) *Mol. Endocrinol.* 14: 27-39. Although these diverse interactions imply a degree of promiscuity, PXR also exhibits specificity. This is illustrated by the fact there are marked differences in the pharmacologic activation profile of PXR across species. For instance, human PXR is activated efficiently by rifampicin and the cholesterol-lowering drug SR12813 (Berkhout et al., (1996) *J. Biol. Chem.* 271: 14376-82; Berkhout et al., (1997) *Atherosclerosis* 133: 203-21), whereas mouse PXR is not (Jones et al., (2000) *Mol. Endocrinol.* 14: 27-39). The mouse version of this receptor is activated by the synthetic steroid pregnenolone 16α-carbonitrile (PCN), whereas the human receptor is not. These data suggest that PXR has evolved in order to detect the different xenobiotic challenges faced by different species. By binding diverse but precise arrays of compounds, PXR exhibits directed promiscuity.

Unraveling the structural basis of how PXR recognizes an array of different endogenous and exogenous compounds, including both small and large ligands, is critical to understanding how harmful compounds are cleared from the body. Such knowledge can also improve our ability to predict and avoid dangerous drug-drug interactions.

Polypeptides, including the ligand binding domain of PXR, have a three-dimensional structure determined by the primary amino acid sequence and the environment surrounding the polypeptide. This three-dimensional structure establishes the polypeptide's activity, stability, binding affinity, binding specificity, and other biochemical attributes. Thus, knowledge of a protein's three-dimensional structure can provide much guidance in designing agents that mimic, inhibit, or improve its biological activity in soluble or membrane bound forms.

The three-dimensional structure of a polypeptide can be determined in a number of ways. Many of the most precise methods employ X-ray crystallography (See, e.g., Van Holde, (1971) Physical Biochemistry, Prentice-

Hall, N. J., 221-39). This technique relies on the ability of crystalline lattices to diffract X-rays or other forms of radiation. Diffraction experiments suitable for determining the three-dimensional structure of macromolecules typically require high-quality crystals. Unfortunately, such crystals have been unavailable for the ligand binding domain of PXR, as well as many other proteins of interest. Thus, high-quality diffracting crystals of the ligand binding domain of PXR would greatly assist in the elucidation of PXR's three-dimensional structure, and would provide insight into the ligand binding properties of PXR.

Clearly, the solved crystal structure of the PXR ligand binding domain would be useful in the design of modulators of activity mediated by all PXR isoforms. Evaluation of the available sequence data has made it clear that PXR shows structural homology with the three-dimensional fold of other proteins.

A solved PXR-ligand crystal structure would provide structural details and insights necessary to design a modulator of PXR that maximizes preferred requirements for any modulator, i.e. potency and specificity. By exploiting the structural details obtained from a PXR-ligand crystal structure, it would be possible to design a PXR modulator that, despite PXR's similarity with other proteins, exploits the unique structural features of PXR. A PXR modulator developed using structure-assisted design would take advantage of heretofore unknown PXR structural considerations and thus be more effective than a modulator developed using homology-based design. Potential or existent homology models cannot provide the necessary degree of specificity. A PXR modulator designed using the structural coordinates of a crystalline form of PXR would also provide a starting point for the development of modulators of other structurally similar proteins.

What is needed, therefore, is a crystallized form of a PXR LBD polypeptide, preferably in complex with a ligand. Acquisition of crystals of the PXR LBD polypeptide will permit the three dimensional structure of the PXR LBD to be determined. Knowledge of this three dimensional structure will facilitate the design of modulators of PXR activity. Such modulators can lead

to therapeutic compounds to treat a wide range of conditions, including conditions associated with toxic xenobiotics.

Summary of the Invention

A substantially pure PXR ligand binding domain polypeptide in crystalline form is disclosed. Preferably, the crystalline form has lattice constants of a = 91.6 Å, b = 91.6 Å, c = 85.0 Å, α = 90°, β = 90°, γ = 90°. More preferably, the crystalline form is a tetragonal crystalline form. Even more preferably, the crystalline form has a space group of $P4_32_12$. Still more preferably, the PXR ligand binding domain polypeptide has the amino acid sequence shown in SEQ ID NO: 4.

In a preferred embodiment, the PXR ligand binding domain polypeptide is in complex with a ligand. More preferably, the ligand is a hypocholesterolemic drug. Even more preferably, the hypocholesterolemic drug is SR12813.

A method for determining the three-dimensional structure of a crystallized PXR ligand binding domain polypeptide to a resolution of about 3.0 Å or better is also disclosed. The method comprises (a) crystallizing a PXR ligand binding domain polypeptide; and (b) analyzing the PXR ligand binding domain polypeptide to determine the three-dimensional structure of the crystallized PXR ligand binding domain polypeptide, whereby the three-dimensional structure of a crystallized PXR ligand binding domain polypeptide is determined to a resolution of about 3.0 Å or better.

A method of designing a modulator of a PXR polypeptide is also disclosed. The method comprises (a) designing a potential modulator of a PXR polypeptide that will form bonds with amino acids in a ligand binding site based upon a crystalline structure of a PXR ligand binding domain polypeptide; (b) synthesizing the modulator; and (c) determining whether the potential modulator modulates the activity of the PXR polypeptide, whereby a modulator of a PXR polypeptide is designed.

In an alternative embodiment, a method of designing a modulator that selectively modulates the activity of a PXR polypeptide in accordance with the

present invention comprises: (a) obtaining a crystalline form of a PXR ligand binding domain polypeptide; (b) evaluating the three-dimensional structure of the crystallized PXR ligand binding domain polypeptide; and (c) synthesizing a potential modulator based on the three-dimensional crystal structure of the crystallized PXR ligand binding domain polypeptide, whereby a modulator that selectively modulates the activity of a PXR polypeptide is designed. Preferably, the method further comprises contacting a PXR ligand binding domain polypeptide with the potential modulator; and assaying the PXR ligand binding domain polypeptide for binding of the potential modulator, for a change in activity of the PXR ligand binding domain polypeptide, or both. More preferably, the crystalline form is such that the three-dimensional structure of the crystallized PXR ligand binding domain polypeptide can be determined to a resolution of about 3.0 Å or better.

In yet another embodiment, a method of designing a modulator of a PXR polypeptide in accordance with the present invention comprises: (a) selecting a candidate PXR ligand; (b) determining which amino acid or amino acids of a PXR polypeptide interact with the ligand using a three-dimensional model of a crystallized protein comprising a PXR LBD; (c) identifying in a biological assay for PXR activity a degree to which the ligand modulates the activity of the PXR polypeptide; (d) selecting a chemical modification of the ligand wherein the interaction between the amino acids of the PXR polypeptide and the ligand is predicted to be modulated by the chemical modification; (e) performing the chemical modification on the ligand to form a modified ligand; (f) contacting the modified ligand with the PXR polypeptide; (g) identifying in a biological assay for PXR activity a degree to which the modified ligand modulates the biological activity of the PXR polypeptide; and (h) comparing the biological activity of the PXR polypeptide in the presence of modified ligand with the biological activity of the PXR polypeptide in the presence of the unmodified ligand, whereby a modulator of a PXR polypeptide is designed. Preferably, the PXR polypeptide is a PXR polypeptide. More preferably, the three-dimensional model of a crystallized protein is a PXR LBD polypeptide with a bound ligand. Even more preferably, the method further

comprises repeating steps (a) through (f), if the biological activity of the PXR polypeptide in the presence of the modified ligand varies from the biological activity of the PXR polypeptide in the presence of the unmodified ligand.

A method for identifying a PXR modulator is also disclosed. The method comprises (a) providing atomic coordinates of a PXR ligand binding domain to a computerized modeling system; and (b) modeling a ligand that fits spatially into a binding cavity or on the surface of the PXR ligand binding domain to thereby identify a PXR modulator. Preferably, the method further comprises identifying in an assay for PXR-mediated activity a modeled ligand that increases or decreases the activity of the PXR.

A method of identifying a PXR modulator that selectively modulates the activity of a PXR polypeptide compared to other polypeptides is disclosed. The method comprises (a) providing atomic coordinates of a PXR ligand binding domain to a computerized modeling system; and (b) modeling a ligand that fits spatially into a binding cavity or on the surface of a PXR ligand binding domain and that interacts with conformationally constrained residues of a PXR that are conserved among PXR isoforms to thereby identify a PXR modulator. Preferably, the method further comprises identifying in a biological assay for PXR–mediated activity a modeled ligand that selectively binds to the PXR ligand binding domain and increases or decreases the activity of the PXR.

An assay method for identifying a compound that inhibits binding of a ligand to a PXR polypeptide is disclosed. The assay method comprises: (a) designing a test inhibitor compound capable of modulating PXR activity, based on the atomic coordinates of a PXR ligand binding domain; (b) synthesizing the test inhibitor compound; (c) incubating a PXR polypeptide with a ligand in the presence of a test inhibitor compound; (d) determining an amount of ligand that is bound to the PXR polypeptide, wherein decreased binding of ligand to the PXR protein in the presence of the test inhibitor compound relative to binding of ligand in the absence of the test inhibitor compound is indicative of inhibition; and (e) identifying the test compound as an inhibitor of ligand binding if decreased ligand binding is observed, whereby

a compound that inhibits binding of a ligand to a PXR polypeptide is identified. Preferably, the ligand is a hypocholesterolemic drug. More preferably, the hypocholesterolemic drug is SR12813.

A method of evaluating a candidate therapeutic agent in humans using mouse model system is disclosed. In a preferred embodiment, the method comprises: (a) providing atomic coordinates of a human PXR ligand binding domain to a computerized modeling system; (b) modeling a candidate therapeutic agent that fits spatially into a binding cavity or on the surface of a human PXR ligand binding domain; (c) providing a mouse PXR polypeptide; (d) selecting one or more mutations to be introduced into an amino acid sequence of the mouse PXR polypeptide, the mutations being selected so as to alter the mouse PXR polypeptide to be similar to a human PXR polypeptide; (e) providing a mutant mouse PXR polypeptide comprising the one or more mutations selected in step (d); (f) contacting the candidate therapeutic agent modeled in step (b) with the mutant mouse PXR polypeptide; (g) determining an effect of the candidate therapeutic agent on the mutant mouse PXR polypeptide; and (h) evaluating the potential of the candidate therapeutic agent for use in humans based on the effect of the candidate therapeutic agent on the mutant mouse PXR polypeptide, whereby the candidate therapeutic agent in humans using a mouse model system is evaluated.

Accordingly, it is an object of the present invention to provide a three dimensional structure of the ligand binding domain of PXR. The object is achieved in whole or in part by the present invention.

An object of the invention having been stated hereinabove, other objects will be evident as the description proceeds, when taken in connection with the accompanying Drawings and Laboratory Examples as best described hereinbelow.

Brief Description of the Drawings

Figure 1 is a ribbon diagram depicting the structure of the ligand binding domain of the human xenobiotic receptor PXR.

Figure 2 is a cut-away view of the ligand binding cavity of the human xenobiotic PXR from a first perspective.

Figure 3 is a cut-away view of the ligand binding cavity of the human xenobiotic PXR from a second perspective.

Figure 4 is an experimentally-observed position of SR12813 in the ligand binding cavity of human PXR. In this figure, intermolecular interactions are shown directly. Amino acid side chains are shown in blue. SR12813 is depicted as a wireframe model and interacting atoms are presented as a spacefilling model comprising light gray spheres.

Figure 5 is an experimentally-observed position of SR12813 in the ligand binding cavity of human PXR. In this figure, intermolecular interactions are shown schematically. van der Waals contacts are indicated by solid arrows and hydrogen bonds with dashed arrows.

Figure 6 is a second experimentally-observed position of SR12813 in the ligand binding cavity of human PXR. In this figure, intermolecular interactions are shown directly. Amino acid side chains are shown in blue. SR12813 is depicted as a wireframe model and interacting atoms are presented as a spacefilling model comprising light gray spheres.

Figure 7 is a second experimentally-observed position of SR12813 in the ligand binding cavity of human PXR. In this figure, intermolecular interactions are shown schematically. van der Waals contacts are indicated by solid arrows and hydrogen bonds with dashed arrows.

Figure 8 is a third experimentally-observed position of SR12813 in the ligand binding cavity of human PXR. In this figure, intermolecular interactions are shown directly. Amino acid side chains are shown in blue. SR12813 is depicted as a wireframe model and interacting atoms are presented as a spacefilling model comprising light gray spheres.

Figure 9 is a third experimentally-observed position of SR12813 in the ligand binding cavity of human PXR. In this figure, intermolecular interactions are shown schematically. van der Waals contacts are indicated by solid arrows and hydrogen bonds with dashed arrows.

Figure 10 is a ribbon diagram depicting two salt bridges (Glu-321 to Arg-410 and Asp-205 to Arg-413) adjacent to the ligand binding cavity of human PXR.

Figures 11A-11D are a series of plots representing the luciferase (normalized luciferase activity counts per second (CPS) x 1000, y axis) and alkaline phosphatase activity of several PXR mutants in the presence of rifampicin or SR12813 (concentration in M, x axis). To generate this plot, CV-1 cells were transfected with expression plasmids for hPXR, D205A, R413A, E321A, or R410A. O=wildtype(WT); Δ=D205A mutant; ■=R413A mutant, in addition to the XREM-CYP3A4-luciferase reporter.

Figure 12 is a computer-generated model depicting the homodimerization interaction observed in human PXR. The various distances between interacting residues are presented in the figure. The individual monomers are denoted by blue and green coloring.

Figure 13 is a ribbon diagram depicting an overview of the crystallographically-observed homodimer of human PXR. Individual monomers are denoted by red and gold coloring, with green and blue coloring representing interacting structure.

Figure 14 is a plot representing the luciferase and alkaline phosphatase activity (normalized luciferase activity counts per second (CPS) x 1000, y axis) of PXR mutants in the presence of rifampicin(RIF) or SR12813(SR) (concentration in M, x axis). To generate this plot, CV-1 cells were transfected with expression plasmids encoding wildtype hPXR or mutant hPXR comprising the mutations W223A, Y225A and the XREM-CYP3A4-luciferase reporter. O=WT-SR; □=W223A, Y225A-RIF; ●=WT-SR; ■= W223A, Y225A-RIF.

Figures 15A-15C are a series of bar graphs demonstrating that four point mutants "humanize" mouse PXR's sensitivity to ligands. CV-1 cells were transfected with expression plasmids for mouse PXR (Fig. 15A), human PXR (Fig. 15B), or R203L, P205S, Q404H, Q407R (mouse→human) mouse PXR (Fig. 15C) and the XREM-CYP3A4-luciferase reporter. On the horizontal axis, bars from left to right represent vehicle, PCN and SR12813.

Brief Description of the Sequences in the Sequence Listing

SEQ ID NO: 1 is a DNA sequence encoding a full-length human PXR polypeptide (Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 2 is an amino acid sequence of a full-length human PXR polypeptide and is derived from the DNA sequence of SEQ ID NO: 1 (Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 3 is a DNA sequence encoding the ligand binding domain of human PXR. The sequence codes for residues 130-434, which corresponds to the ligand binding domain of PXR (Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 4 is an amino acid sequence of the PXR ligand binding domain and is derived from the DNA sequence of SEQ ID NO: 3 (Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 5 is an amino acid sequence of the loop involving residues 309-321 in hPXR (Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 6 is a DNA sequence encoding a full-length human VDR polypeptide (Swiss-Prot Accession No. P11473; GenBank Accession No. J03258).

SEQ ID NO: 7 is an amino acid sequence of a full-lenth human VDR polypeptide and is derived from the DNA sequence of SEQ ID NO: 6 (Swiss-Prot Accession No. P11473; GenBank Accession No. J03258).

SEQ ID NO: 8 is a DNA sequence encoding the ligand binding domain of human VDR. The sequence codes for residues 192-427, which corresponds to the ligand binding domain of PXR (Swiss-Prot Accession No. P11473; GenBank Accession No. J03258).

SEQ ID NO: 9 is an amino acid sequence of the VDR ligand binding domain and is derived from the DNA seqence of SEQ ID NO: 8 (Swiss-Prot Accession No. P11473; GenBank Accession No. J03258).

SEQ ID NO: 10 is a DNA sequence encoding an N-terminal polyhistidine tagged PXR ligand binding domain fusion protein comprising residues 130-434 from human PXR (fragment from Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 11 is an amino acid sequence of the N-terminal polyhistidine tagged fusion protein comprising residues 130-434 from human PXR and is derived from the DNA seqence of SEQ ID NO: 10 (fragment from Swiss-Prot Accession No. O75469; GenBank Accession No. AF061056).

SEQ ID NO: 12 is a DNA sequence encoding residues 623-710 of the human SRC-1 gene (fragment from GenBank Accession No. U59302).

SEQ ID NO: 13 is an amino acid sequence of residues 623-710 of the human SRC-1 gene and is derived from the DNA sequence of SEQ ID NO: 12 (fragment from GenBank Accession No. U59302).

Detailed Description of the Invention

Until disclosure of the present invention presented herein, the ability to obtain crystalline forms of a PXR LBD has not been realized. And until disclosure of the present invention presented herein, a detailed three-dimensional crystal structure of a PXR polypeptide has not been solved.

In addition to providing structural information, crystalline polypeptides provide other advantages. For example, the crystallization process itself further purifies the polypeptide, and satisfies one of the classical criteria for homogeneity. In fact, crystallization frequently provides unparalleled purification quality, removing impurities that are not removed by other purification methods such as HPLC, dialysis, conventional column chromatography, etc. Moreover, crystalline polypeptides are often stable at ambient temperatures and free of protease contamination and other degradation associated with solution storage. Crystalline polypeptides can also be useful as pharmaceutical preparations. Finally, crystallization techniques in general are largely free of problems such as denaturation associated with other stabilization methods (e.g., lyophilization). Once crystallization has been accomplished, crystallographic data provides useful

structural information that can assist the design of compounds that can serve as agonists or antagonists, as described herein below. In addition, the crystal structure provides information useful to map a receptor-binding domain, which could then be mimicked by a small non-peptide molecule that would serve as an antagonist or agonist.

I. Definitions

Following long-standing patent law convention, the terms "a" and "an" mean "one or more" when used in this application, including the claims.

As used herein, the term "mutation" carries its traditional connotation and means a change, inherited, naturally occurring or introduced, in a nucleic acid or polypeptide sequence, and is used in its sense as generally known to those of skill in the art.

As used herein, the term "labeled" means the attachment of a moiety, capable of detection by spectroscopic, radiologic or other methods, to a probe molecule.

As used herein, the term "target cell" refers to a cell, into which it is desired to insert a nucleic acid sequence or polypeptide, or to otherwise effect a modification from conditions known to be standard in the unmodified cell. A nucleic acid sequence introduced into a target cell can be of variable length. Additionally, a nucleic acid sequence can enter a target cell as a component of a plasmid or other vector or as a naked sequence.

As used herein, the term "transcription" means a cellular process involving the interaction of an RNA polymerase with a gene that directs the expression as RNA of the structural information present in the coding sequences of the gene. The process includes, but is not limited to the following steps: (a) the transcription initiation, (b) transcript elongation, (c) transcript splicing, (d) transcript capping, (e) transcript termination, (f) transcript polyadenylation, (g) nuclear export of the transcript, (h) transcript editing, and (i) stabilizing the transcript.

As used herein, the term "expression" generally refers to the cellular processes by which a polypeptide is produced from RNA.

As used herein, the term "transcription factor" means a cytoplasmic or nuclear protein which binds to a gene, or binds to an RNA transcript of a gene, or binds to another protein which binds to a gene or an RNA transcript or another protein which in turn binds to a gene or an RNA transcript, so as to thereby modulate expression of the gene. Such modulation can additionally be achieved by other mechanisms; the essence of a "transcription factor for a gene" pertains to a factor that alters the level of transcription of the gene in some way.

As used herein, the term "hybridization" means the binding of a probe molecule, a molecule to which a detectable moiety has been bound, to a target sample.

As used herein, the term "detecting" means confirming the presence of a target entity by observing the occurrence of a detectable signal, such as a radiologic or spectroscopic signal that will appear exclusively in the presence of the target entity.

As used herein, the term "sequencing" means determining the ordered linear sequence of nucleic acids or amino acids of a DNA or protein target sample, using conventional manual or automated laboratory techniques.

As used herein, the term "isolated" means oligonucleotides substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which they can be associated, such association being either in cellular material or in a synthesis medium. The term can also be applied to polypeptides, in which case the polypeptide will be substantially free of nucleic acids, carbohydrates, lipids and other undesired polypeptides.

As used herein, the term "substantially pure" means that the polynucleotide or polypeptide is substantially free of the sequences and molecules with which it is associated in its natural state, and those molecules used in the isolation procedure. The term "substantially free" means that the sample is at least 50%, preferably at least 70%, more preferably 80% and most preferably 90% free of the materials and compounds with which is it associated in nature.

As used herein, the term "primer" means a sequence comprising two or more deoxyribonucleotides or ribonucleotides, preferably more than three, and more preferably more than eight and most preferably at least about 20 nucleotides of an exonic or intronic region. Such oligonucleotides are preferably between ten and thirty bases in length.

As used herein, the term "DNA segment" means a DNA molecule that has been isolated free of total genomic DNA of a particular species. In a preferred embodiment, a DNA segment encoding a PXR polypeptide refers to a DNA segment that contains SEQ ID NO: 1, but can optionally comprise fewer or additional nucleic acids, yet is isolated away from, or purified free from, total genomic DNA of a source species, such as *Homo sapiens*. Included within the term "DNA segment" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phages, viruses, and the like.

As used herein, the phrase "enhancer-promoter" means a composite unit that contains both enhancer and promoter elements. An enhancer-promoter is operatively linked to a coding sequence that encodes at least one gene product.

As used herein, the phrase "operatively linked" means that an enhancer-promoter is connected to a coding sequence in such a way that the transcription of that coding sequence is controlled and regulated by that enhancer-promoter. Techniques for operatively linking an enhancer-promoter to a coding sequence are well known in the art; the precise orientation and location relative to a coding sequence of interest is dependent, *inter alia*, upon the specific nature of the enhancer-promoter.

As used herein, the terms "candidate substance" and "candidate compound" are used interchangeably and refer to a substance that is believed to interact with another moiety, for example a given ligand that is believed to interact with a complete, or a fragment of, a PXR polypeptide, and which can be subsequently evaluated for such an interaction. Representative candidate substances or compounds include "xenobiotics", such as drugs and other therapeutic agents, carcinogens and environmental pollutants, natural

products and extracts, as well as "endobiotics", such as steroids, fatty acids and prostaglandins. Other examples of candidate compounds that can be investigated using the methods of the present invention include, but are not restricted to, agonists and antagonists of a PXR polypeptide, toxins and venoms, viral epitopes, hormones (e.g., opioid peptides, steroids, etc.), hormone receptors, peptides, enzymes, enzyme substrates, co-factors, lectins, sugars, oligonucleotides or nucleic acids, oligosaccharides, proteins, small molecules and monoclonal antibodies.

As used herein, the term "biological activity" means any observable effect flowing from interaction between a PXR polypeptide and a ligand. Representative, but non-limiting, examples of biological activity in the context of the present invention include dimerization of a PXR and association of a PXR with DNA.

As used herein, the term "modified" means an alteration from an entity's normally occurring state. An entity can be modified by removing discrete chemical units or by adding discrete chemical units. The term "modified" encompasses detectable labels as well as those entities added as aids in purification.

As used herein, the terms "structure coordinates" and "structural coordinates" mean mathematical coordinates derived from mathematical equations related to the patterns obtained on diffraction of a monochromatic beam of X-rays by the atoms (scattering centers) of a molecule in crystal form. The diffraction data are used to calculate an electron density map of the repeating unit of the crystal. The electron density maps are used to establish the positions of the individual atoms within the unit cell of the crystal.

Those of skill in the art understand that a set of structure coordinates determined by X-ray crystallography is not without standard error. For the purpose of this invention, any set of structure coordinates for PXR or a PXR mutant that have a root mean square deviation (RMSD) from ideal of preferably no more than 1.5 Å, more preferably no more than 1.0 Å, and most preferably no more than 0.5 Å when superimposed, using the polypeptide backbone atoms, on the structure coordinates listed in Table 4 shall be

considered identical.

As used herein, the term "space group" means the arrangement of symmetry elements of a crystal.

As used herein, the term "molecular replacement" means a method that involves generating a preliminary model of the wild-type PXR ligand binding domain, or a PXR mutant crystal whose structure coordinates are unknown, by orienting and positioning a molecule whose structure coordinates are known within the unit cell of the unknown crystal so as best to account for the observed diffraction pattern of the unknown crystal. Phases can then be calculated from this model and combined with the observed amplitudes to give an approximate Fourier synthesis of the structure whose coordinates are unknown. This, in turn, can be subject to any of the several forms of refinement to provide a final, accurate structure of the unknown crystal. See, e.g., Lattman, (1985) *Method Enzymol.*, 115: 55-77; Rossmann, ed, (1972) The Molecular Replacement Method, Gordon & Breach, New York.) Using the structure coordinates of the ligand binding domain of hPXR provided by this invention, molecular replacement can be used to determine the structure coordinates of a crystalline mutant or homologue of the hPXR ligand binding domain, or of a different crystal form of the hPXR ligand binding domain.

As used herein, the term "isomorphous replacement" means a method of using heavy atom derivative crystals to obtain the phase information necessary to elucidate the three-dimensional structure of a native crystal (Blundell et al., (1976) Protein Crystallography, Academic Press; Otwinowski, (1991), in Isomorphous Replacement and Anomalous Scattering, (Evans & Leslie, eds.), 80-86, Daresbury Laboratory, Daresbury, United Kingdom). The phrase "heavy-atom derivatization" is synonymous with the term "isomorphous replacement".

As used herein, the terms "β-sheet" and "beta-sheet" mean the conformation of a polypeptide chain stretched into an extended zig-zig conformation. Portions of polypeptide chains that run "parallel" all run in the

same direction. Polypeptide chains that are "antiparallel" run in the opposite direction from the parallel chains.

As used herein, the terms "α-helix" and "alpha-helix" mean the conformation of a polypeptide chain wherein the polypeptide backbone is wound around the long axis of the molecule in a left-handed or right-handed direction, and the R groups of the amino acids protrude outward from the helical backbone, wherein the repeating unit of the structure is a single turnoff the helix, which extends about 0.56 nm along the long axis.

As used herein, the term "unit cell" means a basic parallelepiped shaped block. The entire volume of a crystal can be constructed by regular assembly of such blocks. Each unit cell comprises a complete representation of the unit of pattern, the repetition of which builds up the crystal. Thus, the term "unit cell" means the fundamental portion of a crystal structure that is repeated infinitely by translation in three dimensions. A unit cell is characterized by three vectors a, b, and c, not located in one plane, which form the edges of a parallelepiped. Angles α, β and γ define the angles between the vectors: angle α is the angle between vectors b and c; angle β is the angle between vectors a and c; and angle γ is the angle between vectors a and b. The entire volume of a crystal can be constructed by regular assembly of unit cells; each unit cell comprises a complete representation of the unit of pattern, the repetition of which builds up the crystal.

As used herein, the term "tetragonal unit cell" means a unit cell wherein $a = b \neq c$; and $\alpha = \beta = \gamma = 90°$. The vectors a, b and c describe the unit cell edges and the angles α, β, and γ describe the unit cell angles.

As used herein, the term "crystal lattice" means the array of points defined by the vertices of packed unit cells.

As used herein, the term "ligand binding site" and "ligand binding domain" are used interchangeably and mean that site in a polypeptide where substrate binding occurs. For hPXR, the ligand binding domain comprises the residues 130-434 of the full-length human PXR protein.

As used herein, the term "PXR" means nucleic acids encoding a pregnane X receptor (PXR) nuclear receptor polypeptide that can bind DNA

and/or one or more ligands, and/or has the ability to form multimers. The term "PXR" includes invertebrate homologs; however, preferably, PXR nucleic acids and polypeptides are isolated from vertebrate sources. "PXR" further includes vertebrate homologs of PXR family members, including, but not limited to, mammalian and avian homologs. Representative mammalian homologs of PXR family members include, but are not limited to, murine and human homologs.

As used herein, the terms "PXR gene product", "PXR protein", "PXR polypeptide", and "PXR peptide" are used interchangeably and mean peptides having amino acid sequences which are substantially identical to native amino acid sequences from an organism of interest and which are biologically active in that they comprise all or a part of the amino acid sequence of a PXR polypeptide, or cross-react with antibodies raised against a PXR polypeptide, or retain all or some of the biological activity (e.g., DNA or ligand binding ability and/or dimerization ability) of the native amino acid sequence or protein. Such biological activity can include immunogenicity.

As used herein, the terms "PXR gene product", "PXR protein", "PXR polypeptide", and "PXR peptide" also include analogs of a PXR polypeptide. By "analog" is intended that a DNA or peptide sequence can contain alterations relative to the sequences disclosed herein, yet retain all or some of the biological activity of those sequences. Analogs can be derived from genomic nucleotide sequences as are disclosed herein or from other organisms, or can be created synthetically. Those skilled in the art will appreciate that other analogs, as yet undisclosed or undiscovered, can be used to design and/or construct PXR analogs. There is no need for a "PXR gene product", "PXR protein", "PXR polypeptide", or "PXR peptide" to comprise all or substantially all of the amino acid sequence of a PXR polypeptide gene product. Shorter or longer sequences are anticipated to be of use in the invention; shorter sequences are herein referred to as "segments". Thus, the terms "PXR gene product", "PXR protein", "PXR polypeptide", and "PXR peptide" also include fusion, chimeric or recombinant PXR polypeptides and proteins comprising sequences of the present

invention. Methods of preparing such proteins are disclosed herein and are known in the art.

As used herein, the term "polypeptide" means any polymer comprising any of the 20 protein amino acids, regardless of its size. Although "protein" is often used in reference to relatively large polypeptides, and "peptide" is often used in reference to small polypeptides, usage of these terms in the art overlaps and varies. The term "polypeptide" as used herein refers to peptides, polypeptides and proteins, unless otherwise noted. As used herein, the terms "protein", "polypeptide" and "peptide" are used interchangeably herein when referring to a gene product.

As used herein, the term "modulate" means an increase, decrease, or other alteration of any, or all, chemical and biological activities or properties of a wild-type or mutant PXR polypeptide. The term "modulation" as used herein refers to both upregulation (i.e., activation or stimulation) and downregulation (i.e. inhibition or suppression) of a response.

As used herein, the terms "PXR gene" and "recombinant PXR gene" mean a nucleic acid molecule comprising an open reading frame encoding a PXR polypeptide of the present invention, including both exon and (optionally) intron sequences.

As used herein, the term "gene" is used for simplicity to refer to a functional protein, polypeptide or peptide encoding unit. As will be understood by those in the art, this functional term includes both genomic sequences and cDNA sequences. Preferred embodiments of genomic and cDNA sequences are disclosed herein.

As used herein, the term "DNA sequence encoding a PXR polypeptide" can refer to one or more coding sequences within a particular individual. Moreover, certain differences in nucleotide sequences can exist between individual organisms, which are called alleles. It is possible that such allelic differences might or might not result in differences in amino acid sequence of the encoded polypeptide yet still encode a protein with the same biological activity. As is well known, genes for a particular polypeptide can exist in single or multiple copies within the genome of an individual. Such duplicate

genes can be identical or can have certain modifications, including nucleotide substitutions, additions or deletions, all of which still code for polypeptides having substantially the same activity.

As used herein, the term "intron" means a DNA sequence present in a given gene that is not translated into protein.

As used herein, the term "interact" means detectable interactions between molecules, such as can be detected using, for example, a yeast two hybrid assay. The term "interact" is also meant to include "binding" interactions between molecules. Interactions can, for example, be protein-protein or protein-nucleic acid in nature.

As used herein, the terms "cells," "host cells" or "recombinant host cells" are used interchangeably and mean not only to the particular subject cell, but also to the progeny or potential progeny of such a cell. Because certain modifications can occur in succeeding generations due to either mutation or environmental influences, such progeny might not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

As used herein, the term "agonist" means an agent that supplements or potentiates the bioactivity of a functional PXR gene or protein, of a polypeptide encoded by a gene that is up- or down-regulated by a PXR polypeptide, and/or a polypeptide encoded by a gene that contains a PXR binding site in its promoter region.

As used herein, the term "antagonist" means an agent that decreases or inhibits the bioactivity of a functional PXR gene or protein, or that supplements or potentiates the bioactivity of a naturally occurring or engineered non-functional PXR gene or protein. Alternatively, an antagonist can decrease or inhibit the bioactivity of a functional gene or polypeptide encoded by a gene that is up- or down-regulated by a PXR polypeptide and/or contains a PXR binding site in its promoter region. An antagonist can also supplement or potentiate the bioactivity of a naturally occurring or engineered non-functional gene or polypeptide encoded by a gene that is up- or down-

regulated by a PXR polypeptide, and/or contains a PXR binding site in its promoter region.

As used herein, the terms "chimeric protein" or "fusion protein" are used interchangeably and mean a fusion of a first amino acid sequence encoding a PXR polypeptide with a second amino acid sequence defining a polypeptide domain foreign to, and not homologous with, any domain of one of a PXR polypeptide. A chimeric protein can present a foreign domain that is found in an organism that also expresses the first protein, or it can be an "interspecies" or "intergenic" fusion of protein structures expressed by different kinds of organisms. In general, a fusion protein can be represented by the general formula X— PXR —Y, wherein PXR represents a portion of the protein which is derived from a PXR polypeptide, and X and Y are independently absent or represent amino acid sequences which are not related to a PXR sequence in an organism, which includes naturally occurring mutants. The term "chimeric gene" refers to a nucleic acid construct that encodes a "chimeric protein" or "fusion protein" as defined herein.

As used herein, the term "therapeutic agent" is a chemical entity intended to effectuate a change in an organism. Preferably, but not necessarily, the organism is a human being. It is not necessary that a therapeutic agent be known to effectuate a change in an organism; chemical entities that are suspected, predicted or designed to effectuate a change in an organism are therefore encompassed by the term "therapeutic agent." The effectuated change can be of any kind, observable or unobservable, and can include, for example, a change in the biological activity of a protein.

Representative therapeutic compounds include small molecules, proteins and peptides, oligonucleotides of any length, "xenobiotics", such as drugs and other therapeutic agents, carcinogens and environmental pollutants, natural products and extracts, as well as "endobiotics", such as steroids, fatty acids and prostaglandins. Other examples of therapeutic agents can include, but are not restricted to, agonists and antagonists of a PXR polypeptide, toxins and venoms, viral epitopes, hormones (e.g., opioid peptides, steroids, etc.), hormone receptors, peptides, enzymes, enzyme

substrates, co-factors, lectins, sugars, oligonucleotides or nucleic acids, oligosaccharides, proteins, small molecules and monoclonal antibodies.

II. Description of Tables

Table 1 is a table summarizing the crystal and data statistics obtained from the crystallized ligand binding domain of human PXR. Data on the unit cell are presented, including data on the crystal space group, unit cell dimensions, molecules per asymmetric cell and crystal resolution.

Table 2 is a table comparing human PXR-LBD with nuclear receptor LBDs of known structure.

Table 3 is a table summarizing the calculated buried solvent-accessible surface area of amino acid residues that line the ligand binding cavity of human PXR.

Table 4 is a table of the atomic structure coordinate data obtained from X-ray diffraction from the ligand binding domain of human PXR in complex with a ligand.

Table 5 is a table of the atomic structure coordinate data obtained from X-ray diffraction from human vitamin D receptor that was used in the molecular replacement solution of the human PXR ligand binding domain.

III. General Considerations

Expression of the CYP3A genes is induced at the level of transcription by a variety of xenobiotics, including many that are metabolized by CYP3A. This transcriptional regulation of CYP3A expression provides a mechanism for amplifying the physiologic response during periods of prolonged xenobiotic challenge. The induction of CYP3A expression also represents the basis for an important class of drug-drug interactions. Recently, a member of the nuclear receptor family of ligand-activated transcription facts, the pregnane X receptor (PXR) was shown to serve as a key regulator of CYP3A gene expression. The human ortholog of PXR is alternately referred to as the pregnane-activated receptor (PAR) or the steroid and xenobiotic receptor (SXR). Like other nuclear receptors PXR contains a DNA binding domain

(DBD) and a ligand binding domain (LBD). PXR binds to xenobiotic response elements in the regulatory regions of CYP3A genes as a heterodimer with the 9-cis retinoic receptor (RXR). Notably, PXR is activated by most of the xenobiotics that are known to induce CYP3A gene expression, including commonly used drugs such as the antibiotic rifampicin and the glucocorticoid dexamethasone. Certain endogenous compounds also activate PXR, including the toxic bile acid lithocholic acid and certain C21 steroids (pregnanes). Thus, unlike the steroid, retinoid and thyroid hormone receptors, which are highly selective for their cognate hormone, PXR has evolved in order to detect structurally diverse substrates. Interestingly, there are marked differences in the pharmacologic activation profile of PXR across species. For instance, human PXR is activated efficiently by the rifampicin and the hypocholesterolemic drug SR12813, whereas mouse PXR is not. By contrast, mouse PXR is activated by the synthetic steroid pregnenolone 16a-carbonitrile (PCN), whereas the human receptor is not. These data suggest that PXR has evolved to detect the different xenobiotic challenges faced by different species.

The present invention will usually be applicable *mutatis mutandis* to all PXR polypeptides, as discussed herein based, in part, on the patterns of PXR structure and modulation that have emerged as a consequence of determining the three dimensional structure of human PXR in complex with a ligand. Generally speaking, PXR homologs and orthologs display substantial regions of amino acid homology. Additionally, the PXRs display an overall structural motif comprising three modular domains:

1) a variable amino-terminal domain;

2) a highly conserved DNA-binding domain (DBD); and

3) a less conserved carboxy-terminal ligand binding domain (LBD).

The modularity of PXR permits different domains of each protein to separately accomplish different functions, although the domains can influence each other. The separate function of a domain is usually preserved when a particular domain is isolated from the remainder of the protein. Using conventional protein chemistry techniques, a modular domain can sometimes

be separated from the parent protein. Using conventional molecular biology techniques, each domain can usually be separately expressed with its original function intact or, as discussed herein below, chimeric proteins comprising two different proteins can be constructed, wherein the chimeric proteins retain the properties of the individual functional domains of the respective polypeptides from which the chimeric proteins were generated.

The amino terminal domain of PXR is the least conserved of the three domains. This domain is involved in transcriptional activation and, in some cases, its uniqueness can dictate selective receptor-DNA binding and activation of target genes by PXR.

The DBD is the most conserved structure in PXR. It typically contains about 70 amino acids that fold into two zinc finger motifs, wherein a zinc ion coordinates four cysteines. The DBD generally contains two perpendicularly oriented α-helices that extend from the base of the first and second zinc fingers. The two zinc fingers function in concert along with non-zinc finger residues to direct the PXRs to specific target sites on DNA. Various amino acids in the DBD influence spacing between two half-sites (which usually comprises six nucleotides) for receptor homodimerization. The optimal spacings facilitate cooperative interactions between DBDs, and D box residues are part of the dimerization interface. Other regions of the DBD facilitate DNA-protein and protein-protein interactions required for PXR-RXR heterodimerization.

The LBD is the second most highly conserved domain in these receptors. Whereas the integrity of several different LBD sub-domains is important for ligand binding, truncated molecules containing only the LBD can retain normal ligand binding activity. This domain also participates in other functions, including dimerization, nuclear translocation and transcriptional regulation activities. Importantly, this domain can bind a ligand and can undergo ligand-induced conformational changes. Ligand binding allows the activation domain to serve as an interaction site for essential co-activator proteins that function to stimulate or inhibit transcription.

The carboxy-terminal activation subdomain is in close three-dimensional proximity in the LBD to the ligand, so as to allow for ligands bound to the LBD to coordinate (or interact) with amino acid(s) in the activation subdomain. As disclosed herein, the LBD of PXR is expressed, crystallized and its three dimensional structure determined. Computational and other methods for the design of ligands to the LBD are also disclosed.

IV. Production of PXR Polypeptides

The native and mutated PXR polypeptides, and fragments thereof, of the present invention can be chemically synthesized in whole or part using techniques that are well-known in the art (See, e.g., Creighton, (1983) Proteins: Structures and Molecular Principles, W.H. Freeman & Co., New York, incorporated herein in its entirety). Alternatively, methods that are well known to those skilled in the art can be used to construct expression vectors containing a partial or the entire native or mutated PXR polypeptide coding sequence and appropriate transcriptional/translational control signals. These methods include *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* recombination/genetic recombination. See, for example, the techniques described in Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York, and Ausubel et al., (1989) Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, New York, both incorporated herein in their entirety.

A variety of host-expression vector systems can be utilized to express a PXR coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing a PXR coding sequence; yeast transformed with recombinant yeast expression vectors containing a PXR coding sequence; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing a PXR coding sequence; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed

with recombinant plasmid expression vectors (e.g., Ti plasmid) containing a PXR coding sequence; or animal cell systems. The expression elements of these systems vary in their strength and specificities.

Depending on the host/vector system utilized, any of a number of suitable transcription and translation elements, including constitutive and inducible promoters, can be used in the expression vector. For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage λ, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like can be used. When cloning in insect cell systems, promoters such as the baculovirus polyhedrin promoter can be used. When cloning in plant cell systems, promoters derived from the genome of plant cells, such as heat shock promoters; the promoter for the small subunit of RUBISCO; the promoter for the chlorophyll a/b binding protein) or from plant viruses (e.g., the 35S RNA promoter of CaMV; the coat protein promoter of TMV) can be used. When cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter) can be used. When generating cell lines that contain multiple copies of the tyrosine kinase domain DNA, SV40-, BPV- and EBV-based vectors can be used with an appropriate selectable marker.

V. Formation of PXR Ligand Binding Domain Crystals

In one embodiment, the present invention provides crystals of PXR. The crystals were obtained using the methodology disclosed in the Laboratory Examples. The PXR crystals, which can be native crystals, derivative crystals or co-crystals, have tetragonal unit cells (a tetragonal unit cell is a unit cell wherein a = b ≠ c, and wherein $\alpha = \beta = \gamma = 90°$) and space group symmetry $P4_32_12$. There is one PXR molecule in the asymmetric unit. In the PXR crystalline form, the unit cell has dimensions of a= b= 91.6 Å, c= 85.0 Å, and $\alpha = \beta = \gamma = 90°$.

The structures of the apo and SR12813-bound forms of PXR were solved by molecular replacement using the structure of the human vitamin D

receptor as a search model (Rochel et al., (2000) *Mol. Cell* 5: 173-79, PDB ID: 1DB1; GenBank Accession No. XM007046; available online at http://www.rcsb.org/pdb/). The apo form was refined to a resolution of about 2.5 Å. The ligand-bound form was refined to a resolution of about 2.75 Å.

The heavy atom derivatized form of the PXR LBD-ligand structure can be solved using single isomorphous replacement anomalous scattering (SIRAS) techniques and/or multiwavelength anomalous diffraction (MAD) techniques. In the SIRAS method of solving protein crystals, a derivative crystal is prepared that contains an atom that is heavier than the other atoms of the sample. Heavy atom derivative crystals are commonly prepared by soaking a crystal in a solution containing a selected heavy atom salt. For example, some heavy atom derivative crystals have been prepared by soaking a crystalline form of the protein of interest in a solution of methyl mercury chloride (MeHgCl).

Another representative heavy atom that can be incorporated into a derivative crystal is iodine. Heavy atoms can associate with the protein of interest, or can be localized in a ligand that associates with a protein of interest. In the present invention, the latter approach was taken. Specifically, an iodine-containing form of the ligand SR12813 was co-crystallized with the PXR LBD and was used to assist in the positioning of the ligand relative to the PXR LBD.

Analysis of derivative crystals takes advantage of differences in the reflections from the derivative crystal as compared to the underivatized crystal. Symmetry-related reflections in the X-ray diffraction pattern, which are usually identical, are altered by the anomalous scattering contribution of the heavy atoms. The measured differences in symmetry-related reflections are used to determine the position of the heavy atoms, leading to an initial estimation of the diffraction phases, and subsequently, an electron density map is prepared. The prepared electron density map is then used to identify the position of the other atoms in the sample.

V.A. Preparation of PXR Crystals

The native and derivative co-crystals, and fragments thereof, disclosed in the present invention can be obtained by a variety of techniques, including batch, liquid bridge, dialysis, vapor diffusion and hanging drop methods (See, e.g., McPherson, (1982) Preparation and Analysis of Protein Crystals, John Wiley, New York.; McPherson, (1990) *Eur. J. Biochem.* 189:1-23.; Weber, (1991) *Adv. Protein Chem.* 41:1-36). In a preferred embodiment, the vapor diffusion and hanging drop methods are used for the crystallization of PXR polypeptides and fragments thereof.

In general, native crystals of the present invention are grown by dissolving substantially pure PXR LBD polypeptide or a fragment thereof in an aqueous buffer containing a precipitant at a concentration just below that necessary to precipitate the protein. Water is removed by controlled evaporation to produce precipitating conditions, which are maintained until crystal growth ceases.

In a preferred embodiment of the invention, native crystals are grown by vapor diffusion (See, e.g., McPherson, (1982) Preparation and Analysis of Protein Crystals, John Wiley, New York; McPherson, (1990) *Eur. J. Biochem.* 189:1-23). In this method, the polypeptide/precipitant solution is allowed to equilibrate in a closed container with a larger aqueous reservoir having a precipitant concentration optimal for producing crystals. Generally, less than about 25 μL of PXR LBD polypeptide solution is mixed with an equal volume of reservoir solution, giving a precipitant concentration about half that required for crystallization. This solution is suspended as a droplet underneath a coverslip, which is sealed onto the top of the reservoir. The sealed container is allowed to stand, until crystals grow. Crystals generally form within two to six weeks, and are suitable for data collection within approximately seven to ten weeks. Of course, those of skill in the art will recognize that the above-described crystallization procedures and conditions can be varied.

V.B. Preparation of Derivative Crystals

Derivative crystals of the present invention, e.g. heavy atom derivative crystals, can be obtained by soaking native crystals in mother liquor containing salts of heavy metal atoms. Alternatively, a ligand comprising a heavy atom can be associated with a protein, and subsequently co-crystallized. Such derivative crystals are useful for phase analysis in the solution of crystals of the present invention. This mechanism provides derivative crystals suitable for use as isomorphous replacements in determining the X-ray crystal structure of a PXR polypeptide. Additional reagents useful for the preparation of the derivative crystals of the present invention will be apparent to those of skill in the art after review of the disclosure of the present invention presented herein.

V.C. Preparation of Co-crystals

Co-crystals of the present invention can be obtained by soaking a native crystal in mother liquor containing compounds known or predicted to bind the LBD of a PXR, or a fragment thereof. Alternatively, co-crystals can be obtained by co-crystallizing a PXR LBD polypeptide or a fragment thereof in the presence of one or more compounds known or predicted to bind the polypeptide. In a preferred embodiment of the present invention, for example, the ligand SR12831, or an iodinated form of the ligand SR12831, is co-crystallized with a PXR-SRC-1 complex.

V.D. Solving a Crystal Structure of the Present Invention

Crystal structures of the present invention can be solved using a variety of techniques including, but not limited to, isomorphous replacement anomalous scattering or molecular replacement methods. Computer software packages will also be helpful in solving a crystal structure of the present invention. Applicable software packages include but are not limited to X-PLOR™ program (Brünger, (1992) *X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, Connecticut; X-PLOR is available from Molecular Simulations, Inc., San Diego, California),

Xtal View (McRee, (1992) *J. Mol. Graphics* 10: 44-47; X-tal View is available from the San Diego Supercomputer Center), SHELXS 97 (Sheldrick (1990) *Acta Cryst.* A46: 467; SHELX 97 is available from the Institute of Inorganic Chemistry, Georg-August-Universität, Göttingen, Germany), HEAVY (Terwilliger, Los Alamos National Laboratory) can be used and SHAKE-AND-BAKE (Hauptman, (1997) *Curr. Opin. Struct. Biol.* 7: 672-80; Weeks et al., (1993) *Acta Cryst.* D49: 179; available from the Hauptman-Woodward Medical Research Institute, Buffalo, New York). See also, Ducruix & Geige, (1992) Crystallization of Nucleic Acids and Proteins: A Practical Approach, IRL Press, Oxford, England, and references cited therein.

VI. Summary of Results for the PXR Ligand binding Domain

Two structures of the LBD of human PXR are disclosed in the present invention. First, a 2.5 Å apo structure and second, a 2.75 Å structure in complex with the high affinity ligand SR12813. More specifically, a 1:1 complex, comprising residues 130-434 of human PXR and residues 623-710 of the human transcriptional co-activator protein SRC-1, was incubated in the crystallization mixture. The mixture was crystallized. The 88-amino acid co-activator peptide SRC-1 was not observed in the crystal structure, and the examination of extensively washed and dissolved crystals by SDS-PAGE revealed that this co-activator fragment was not present in the crystals of the PXR LBD. Residues 142-177 and 198-431 of human PXR LBD were traced in the structures reported in an aspect of the present invention, while ordered electron density of the remaining amino acids, including the 20-residue stretch from 178-197, was not observed at any time during refinement.

The hydrophobic ligand binding cavity of human PXR contains a small number of polar residues, permitting the ligand SR12813 to bind in three distinct orientations. Despite the promiscuity of this receptor, human and mouse PXR are activated by different collections of xenobiotics. Critical residues responsible for this directed promiscuity are disclosed in the present invention. Interestingly, the mutation of only four residues "humanizes" murine PXR. That is, the mutation of four residues is observed to change the

binding profile of murine PXR from a murinic profile to that of the human PXR. The crystal structures of the present invention disclose important insights into how human PXR detects xenobiotics and thus, can be useful in predicting and avoiding drug-drug interactions.

VI.A. Overall Structure of the Human PXR Ligand Binding Domain

Two structures of the ligand binding domain (LBD) of hPXR were determined by molecular replacement (Navaza & Saludjian, (1997) *Method Enzymol.* 276A: 581-94) and were refined with CNS (Brünger et al., (1998) *Acta Crystallogr. D* 54: 905-21). As noted herein above, the solved structures disclosed in the present invention include the human PXR LBD apo structure, which was solved to a resolution of 2.5 Å, and the structure of human PXR in complex with the high-affinity (K_d = 41 nM) ligand SR12813, which was solved to a resolution of 2.75 Å. The apo structure was refined to final R and R_{free} values of 0.209 and 0.284 respectively. The ligand-bound structure (PXR-SR12813) was refined to R and R_{free} values of 0.213 and 0.274, respectively. These and other calculated and observed crystallographic data are summarized in Table 1. Residues 142-177 and 198-431 of the human PXR LBD were traced in the structures reported in the present invention.

The human PXR LBD is an "α-helical sandwich" comprising three layers: α1/α3, α7/α10, and a middle layer comprising α-helices 4, 5 and 8. See Figures 1 and 13. This region of the molecule is similar to nuclear receptor ligand binding domains having known structure (Moras & Gronenmeyer, (1998) *Curr. Opin. Cell Biol.* 10: 384-91; Weatherman et al., (1999) *Annu. Rev. Biochem.* 68: 559-81). The standard three-stranded β-sheet is expanded to a five-stranded antiparallel β-sheet in the human PXR LBD, comprising β-strands 1, 1', 2, 3 and 4, and the ligand binding cavity is localized at the bottom of the molecule. This structural feature is shown in Figure 1. The structures of the apo and ligand-bound forms of the hPXR LBD are essentially identical, exhibiting a root-mean-square deviation (RMSD) of 0.68 Å over the Cα positions, and 0.89 Å over all atoms. In both the apo and SR12813-bound structures, the activation function 2 helix (α-AF2), which

plays a critical role in transcriptional activation by nuclear receptors, is packed against the body of the receptor in a position that appears permissive for coactivator interactions.

The human PXR (SEQ ID NOs: 1 and 2) LBD (SEQ ID NOs: 3 and 4) is most closely related in structure to the vitamin D receptor (VDR) (Rochel et al., (2000) *Mol. Cell* 5: 173-79), with which it shares 45% sequence identity (SEQ ID NOs: 8 and 9) and exhibits a 1.8 Å RMSD over 225 equivalent Cα positions. Table 2 illustrates this similarity. More limited structural similarity, between 2.4 and 2.9 Å RMSD, is shared with the ligand binding domains of the retinoid x receptor (RXR), peroxisome proliferator-activated receptor γ (PPARγ), and the estrogen and progesterone receptors (Bourget et al., (1995) *Nature* 375: 377-82; Renaud et al., (1995) *Nature* 378: 681-89; Nolte et al., (1998) *Nature* 395: 137-43; Williams & Sigler, (1998) *Nature* 393: 392-96; Shiau et al., (1998) *Cell* 95: 927-37; Brzozowski et al., (1997) *Nature* 389: 753-58). This observation is also presented in Table 2.

VI.B. Novel Structural Features of the Human PXR LBD

While human PXR is similar to known nuclear receptor ligand binding domains, it contains several distinct features that appear critical to its function as a promiscuous xenobiotic receptor. First, the variable region typically found between α1 and α3 in other nuclear receptors is replaced by a four-residue turn in human PXR, as depicted in Figure 1. In the peroxisome proliferator-activated receptors (PPARs), this region contains α2 and is the proposed ligand access site for the binding cavity (Nolte et al., (1998) *Nature* 395: 137-43; Gampe et al., (2000) *Mol. Cell* 5: 545-55. This region cannot be the access site in hPXR because it is occluded by the turn formed by residues 171-174. These structural features suggest that a distinct ligand entry site exists in hPXR.

Second, helix α6, which is common to many nuclear receptors, is replaced in hPXR by a conserved, flexible loop composed of residues 309-321 (See Figures 1 and 13). This region lies adjacent in space to the ligand binding cavity of hPXR. This feature might be related to the ability of the

hPXR receptor to accommodate both small and large ligands in the binding cavity.

Third, hPXR has two additional β-strands that have not previously been observed in a nuclear receptor ligand binding domain. This feature is depicted in Figure 1. Weatherman et al., (1999) *Annu. Rev. Biochem.* 68: 559-81. These additional strands form the fourth (β1, residues 210 to 217) and fifth (β1', residues 221 to 226) strands of a five-stranded anti-parallel β-sheet. An "insertion domain" containing roughly the same number of residues was engineered out of the VDR prior to crystallization and structure determination (Rochel et al., (2000) *Mol. Cell* 5:173-79). However, small angle x-ray scattering (Rochel et al., *Eur. J. Biochem.* 268: 971-79) suggests that the 51-amino acid insertion domain in VDR (residues 164-216) is unstructured, and shares only 12% sequence identity with the analogous 45-amino acid (residues 184 to 229) in PXR. β1', at the base of this five-stranded beta-sheet, forms a novel homodimerization interaction in hPXR. As discussed herein below, these structural features appear to play a role in ligand binding, accessibility of the binding cavity, and the formation of higher-order oligomers of hPXR.

VI.C. The Ligand Binding Cavity of Human PXR

The ligand binding cavity of the hPXR-LBD is largely hydrophobic and might be flexible in nature, in order to accommodate both small and large ligands. The 28 amino acid residues desolvated by the binding of SR12813, which are presented in Table 3, are considered to be lining the cavity. The structures of the apo and the ligand bound cavities are similar, exhibiting a 1.12 Å RMSD over all atoms in these 28 residues. Table 2 shows that the binding cavity volume of 1,150 $Å^3$, making it larger than most known nuclear receptor ligand binding cavities. Twenty of the cavity-lining residues are hydrophobic, four are polar (Ser-208, Ser-247, Cys-284 and Gln-285), and four are charged or potentially charged (Glu-321, His-327, His-407 and Arg-410). This observation is depicted via graphics in Figures 2 and 3 and is additionally presented in Table 3. As described further herein below, Glu-321

and Arg-410 are involved in a salt bridge, effectively neutralizing their charged character adjacent to the ligand binding cavity. Thus, an electrostatic view of the inner surface of this ligand binding cavity reveals a relatively uncharged and hydrophobic environment, as seen in Figures 2 and 3.

Spaced throughout this cavity are five critical polar residues which form key binding interactions with SR12813 (Ser-208, Ser-247, Gln-285, His-407 and Arg-410, also indicated in Figures 2 and 3). These residues are capable of hydrogen bonding and are spaced evenly throughout the upper portion of the ligand binding cavity. The precise placement of polar residues is found to be critical for defining the activation profile of mouse PXR with respect to specific ligands.

hPXR is able to bind both small and large ligands. The loop involving residues 309-321 in hPXR (comprising the sequence **EDTAGGFQQLLL**E, SEQ ID NO: 5) replaces α6 observed previously in reported nuclear receptor ligand binding domain structures. This highly mobile region spans the space between the C-terminus of β4 to the N-terminus of α7, and exhibits a mean thermal displacement parameter of 82.3 $Å^2$ over main-chain atoms despite persistent electron density. Nine of the thirteen residues in this loop are completely conserved in the mammalian PXR molecules of known sequence (highlighted in bold in SEQ ID NO: 5 above, and depicted in Figures 2 and 3), including the solvent-exposed hydrophobic residues of SEQ ID NO: 5, Phe-315, Leu-318, Leu-319 and Leu-320. The 309-321 loop is linked to the ligand binding cavity of hPXR by a non-solvent accessible pore. This observation is depicted in Figures 2 and 3. The binding of a particularly large and hydrophobic ligand in the binding cavity (e.g. rifampicin) might force the pore to this loop region to open, enlarging the ligand binding region of hPXR and lining it with additional hydrophobic amino acid side chains.

VI.D. Three Orientations of SR12813 in the Ligand binding Cavity

Three distinct binding modes of the high-affinity ligand SR12813 (positions 1, 2 and 3) were observed in the ligand binding cavity of hPXR, shown in Figures 4-9. These orientations were identified during structural

refinement and were rigorously confirmed using difference maps involving data obtained from crystals containing an iodinated form of SR12813 (I-SR12813 in Figures 4-9). As shown in Figures 4-9, each orientation forms distinct interactions with residues that line the ligand binding cavity of PXR. While ligand position 3 forms the most hydrophilic interactions with the protein, positions 1 and 2 were clearly indicated in the detailed examination of difference maps and in refinement. Indeed, when any one of the orientations is not considered in an unbiased refinement, clear positive electron density appears in difference maps to indicate its presence in the binding cavity of the protein.

Of the nineteen residues involved in contacting these orientations of SR12813, only Phe-288 interacts with all three ligand conformations. A phenylalanine residue is conserved at position 288 in the known mammalian PXR sequences. The remainder of the hydrophobic residues contact either two orientations (Leu-209, Phe-251, Phe-281, Trp-299) or one orientation (Leu-206, Val-211, Leu-240, Met-243, Met-246, Met-323, Leu-324, Phe-420) of the ligand. Among the polar side chains, two (Ser-247, His-407) interact with two orientations of the ligand, while the remaining four (Ser-208, Cys-284, Gln-285, Arg-410) interact with only one orientation. All but five of these residues are conserved by either identity or similarity in the mammalian PXR sequences reported to date, an observation presented in Table 3.

VI.D.1. *The Ligand Binding Promiscuity of PXR*

The nuclear receptor PXR serves as a key component of the body's defense mechanism against xenobiotics by detecting these compounds and regulating the transcription of genes involved in their metabolism (Waxman, (1999) *Arch. Biochem Biophys.* 369: 11-23; Savas et al., (1999) *Mol. Pharmacol.* 56: 851-57). Unlike other nuclear receptors, which evolved to interact selectively with their cognate hormone, PXR evolved the ability to interact promiscuously with a structurally diverse collection of hydrophobic compounds. The x-ray crystal structures of the apo and SR12813-bound

PXR-LBD reveal several important features underlying PXR's promiscuous ligand binding properties.

First, the volume of the PXR ligand binding cavity is >1,100 $Å^3$, which is substantially larger than that of many other nuclear receptors including the progesterone, estrogen, retinoid, and thyroid hormone receptors, an observation represented in Table 2. The PXR structure disclosed in the present invention also suggests that the ligand binding cavity may be capable of expanding to an even larger size. A stretch of conserved residues 309-321 (identified in Figures 1-3) loop out and away from the ligand binding cavity of PXR in the structures reported here, exposing conserved hydrophobic residues to solvent. This loop is connected to the existing ligand binding cavity by a solvent-inaccessible pore. These observations suggest that PXR might rearrange its ligand binding cavity when confronted by larger ligands, such as rifampicin. If the pore were to open between the ligand binding cavity and the 309-321 loop, the conserved hydrophobic residues currently exposed to solvent might be used to line a larger hydrophobic binding cavity. Thus, structural flexibility in the ligand binding cavity can contribute to PXR's ability to bind both to small hydrophobic ligands as well as to the much larger compound rifampicin.

Second, the position of the ligand entry site appears to be unique in PXR. Instead of being near the AF2-helix as in the steroid, retinoid, and thyroid hormone receptors or in the front of the molecule as in the PPARs, the entry site appears to be located on the back of the molecule between α7 and α10. The putative entrance-exit path to the ligand binding cavity in hPXR might be gated by two salt bridges (i.e. between Asp-205 and Arg-413, and Glu-321 and Arg-410), as shown in Figure 10. Mutagenesis studies disclosed in the present invention show that Arg-410 and Asp-205, which are in van der Waals contact with each other, are critical to defining the appropriate basal activity level for PXR. When Asp-205 is mutated to alanine, the basal activity level of PXR drops dramatically in cell-based reporter gene assays. In contrast, changing Arg-410 to alanine increases the basal activity level of PXR activity, as indicated in Figure 11. The interactions of Glu-321 and Arg-

413 with Arg-410 and Asp-205, respectively, might be necessary simply to define the locations of Asp-205 and Arg-410; the correct positioning of these two residues appears to tune the transcriptional activity of PXR. Perhaps during the opening and closing of this gate to allow for the movement of ligands, a vital intermediate is formed by the interaction of Arg-410 with Asp-205. The drop in basal activity level in the Asp-205-Ala mutant might also indicate that this form of the protein is less stable than the wildtype. Asp-205 could be critical for ordering the 198-210 region of the molecule; the loss of this residue could introduce flexibility to PXR that affects its stability and basal activation level.

Third, the ligand binding cavity of PXR is relatively smooth and hydrophobic, but contains a small number of key polar residues. Twenty of the 28 residues lining the binding cavity are hydrophobic; the remaining polar residues are spaced roughly evenly throughout the cavity, offering the potential of a small number of hydrogen bonds. The character of the PXR ligand cavity mirrors the character of the majority of the ligands known to activate PXR, such as phenobarbital, clotrimazole, RU486, hyperforin, PCN and lovastatin, which are largely hydrophobic and uncharged, but which contain oneto four functional groups capable of forming hydrogen bonds.

A combination of hydrophobic and shuffled polar interactions allows the potent PXR agonist SR12813 to bind to the hPXR ligand binding cavity in three distinct orientations. Although SR12813 occupies essentially the same portion of the cavity in each of the three binding modes, each mode is facilitated by a different set of hydrogen bonding and hydrophobic interactions (Figures 4-6). For example, Ser-247 forms a hydrogen bond with the phenolic hydroxyl group on SR12813 in one binding mode, and a hydrogen bond with a phosphate group in another. Thus, PXR does not define a specific ligand binding surface that will drive the binding of ligands in one mode. Instead, PXR offers an essentially smooth and uncharged ligand binding surface, with evenly spaced hydrogen bond donors and acceptors (Figures 2 and 3). Such an arrangement not only permits PXR to bind to structurally distinct ligands, but also allows single ligands to bind in multiple modes. This binding mode

stands in sharp contrast to other nuclear receptor-ligand interactions, which have evolved to be highly specific.

VI.E. Electrostatic Interactions Adjacent to the Ligand binding Cavity

Two salt-bridges occur across the region of the ligand binding cavity that is closest to the surface of hPXR and are depicted in Figure 10: Arg-410 with Glu-321, and Arg-413 with Asp-205. Asp-205, Glu-321 and Arg-413 are almost completely conserved in the PXR molecules of known sequence; rabbit PXR contains a glutamic acid at position 205. Arg-410, however, is less conserved; in human and rabbit PXR, this residue is an arginine or lysine, while in rat and mouse it is a glutamine.

To examine the functional relevance of the Asp-205/Arg-413 and Glu-321/Arg-410 salt bridges, each of these residues was individually mutated to alanine. Expression plasmids for each of the four mutant PXRs were co-transfected into CV-1 cells together with the XREM-CYP3A4-LUC reporter plasmid containing the distal enhancer and proximal promoter of CYP3A4, which include two high-affinity PXR binding sites, driving expression of the luciferase gene (Goodwin et al., (1999) *Mol. Pharmacol.* 56: 1329-39). Transfected cells were treated with increasing concentrations of either SR12813 or rifampicin, as illustrated in the plots of Figure 11. Mutation of Asp-205 resulted in a marked decrease in the basal (ligand-independent) transcriptional activity of PXR, while mutation of Arg-413 resulted in a more modest reduction in PXR basal activity. In the other salt bridge, mutation of Glu-321 resulted in a marked increase in the basal activity of PXR; conversely, the Arg-410-Ala mutant had reduced basal activity. Taken together, these data demonstrate that the residues that form these salt bridges are important in determining the basal activity of the PXR. Moreover, the data highlight the asymmetry in the functions of the residues comprising the salt bridges: mutation of either Arg-410 or Asp-205 has a more dramatic effect than alteration of their salt bridge partners, suggesting that these two residues are particularly critical for determining the basal transcriptional activity of PXR.

The mutant PXR LBD polypeptides Asp-205-Ala, Arg-413-Ala and Glu-321-Ala were all activated by both SR12813 and rifampicin, although the dynamic range of the response was compressed relative to the wildtype receptor for the Glu-321-Ala mutant. Interestingly, rifampicin was a more potent activator of the Asp-205-Ala mutant than wildtype PXR. This effect was not seen with SR12813, suggesting that SR12813 and rifampicin bind to PXR in different ways. These results also indicate that mutation of residues involved in salt bridges can affect the potency and selectivity with which compounds activate PXR.

VI.F. hPXR Forms a Novel Homodimer

A remarkable observation arising from these structural studies was that the ligand binding domain of hPXR forms a homodimer through the β-sheet structure that is unique to PXR. β1' from the β1-β1' region of this structure forms an ideal anti-parallel β-sheet interaction with β1' from a crystallographically-related monomer, as shown in Figure 12. The dimer interface appears to be stabilized in two ways. First, six 2.9–3.0 Å main-chain to main-chain hydrogen bonds in an anti-parallel β-sheet pattern occur across the interface. This generates a ten-stranded anti-parallel β-sheet between the two monomers, highlighted in Figure 13. Second, two aromatic side chains are observed to be buried which lends a large degree of specificity to this interaction. Notably, the Trp-223 side chains from each monomer are locked across the dimer interface, and thus are completely protected from solvent, permitting them to form an "offset-edge" stacking interaction frequently observed between tryptophans (Samanta et al., (1999) *Acta Crystalog. D* 55: 1421-27). Tyr-225 is also involved in this interface and becomes buried in the dimer, packing against Pro-175 from the region leading up to residue 177. The position of Trp-223 is further stabilized by a 3.0 Å hydrogen bond between the main-chain carbonyl oxygen of Pro-175 and the indol ring nitrogen of the tryptophan side chain. In all, this highly specific dimer interface buries about 1,600 $Å^2$ of solvent accessible surface area on each monomer. Trp-223, Tyr-225 and Pro-175 are conserved in all the reported

mammalian PXR sequences. The evolutionary conservation of the residues involved in this interaction supports its role in the biological function of PXR.

To assess whether this homodimerization interface is functionally relevant *in vivo*, reporter gene experiments were performed with a PXR double mutant in which Trp-223 and Tyr-225 were mutated to alanine. The Trp-223-Ala/Tyr-225-Ala mutant hPXR polypeptide had dramatically reduced basal activity and was only very weakly responsive to either SR12813 or rifampicin, as shown in Figure 14. Gel mobility shift assays revealed that this double mutant retains the ability to bind to DNA as a heterodimer with RXRα. Taken together, these data suggest an unprecedented role for a novel homodimerization interface in PXR-mediated transcriptional activation.

A model for a putative heterotetrameric complex formed by the ligand binding domains of hPXR and RXRα was generated by superimposing PXR on PPARγ in the heterodimeric structure of PPARγ and RXRα (Gampe et al., (2000) *Mol Cell* 5: 545-55). The interface between PXR and RXRα generated in this fashion is nearly ideal, requiring no manual optimization and producing no van der Waals overlaps between atoms. Importantly, the homodimerization of PXR does not interfere with the PXR/RXRα interface in the heterotetramer, suggesting that a heterotetramer of PXR and RXRα can or could occur *in vivo*.

VI.G Structure-Based Receptor Design: Altering the Ligand Specificity of Mouse PXR By Four Targeted Site-Directed Mutants

Marked differences in the pharmacological activation profiles of PXR occur across species. For example, human and rabbit PXR are activated efficiently by SR12813 whereas mouse PXR is not. Conversely, mouse PXR is activated more efficiently by the synthetic steroid PCN than either the human or rabbit orthologs (Jones et al., (2000) *Mol. Endocrinol.* 14: 27-39; Savas et al., (2000) *Drug Metab Dispos.* 28: 529-37). To verify the importance of polar residues in the selective PXR-SR12813 interaction,

mouse PXR was "humanized" – that is, its selectivity was altered such that it would respond to SR12813 but not to PCN.

Four residues that interact with SR12813 in the hPXR crystal structure and differ between human and mouse PXR were chosen for mutagenesis: Leu-206/Arg-203, Ser-208/Pro-205, His-407/Gln-404, Arg-410/Gln-407 (human residues in each pair are listed first). Each of these residues was mutated in the context of a mouse PXR expression plasmid to the corresponding amino acid in the hPXR (R203L, P205S, Q404H, Q407R; using the mouse numbering scheme). As expected, wildtype mouse PXR responded to PCN and was only weakly activated by SR12813 in reporter assays (Figure 5). However, the R203L, P205S, Q404H, Q407R mutant (mouse→human PXR) was no longer activated by PCN but was activated efficiently by SR12813, mirroring the hPXR activation profile, as shown by the bar graphs presented in Figure 15. Thus, mouse PXR's response to particular ligands was "humanized" by employing the crystal structure to design targeted mutants. These results show that the selectivity of PXR for certain ligands resides in the number and position of only a few polar residues within the molecule's large, hydrophobic ligand binding cavity.

VI.G.1. Species-Specific Activation by Ligand in the PXRs

PXR is promiscuous in that it binds a wide variety of endogenous compounds and xenobiotics. However, this receptor also exhibits specificity as demonstrated by the fact that PXRs from different species show distinct activation profiles. To test the hypothesis that key polar residues determine specificity, the crystal structures of hPXR disclosed in the present invention were employed to design an altered form of PXR. Mutation of only four residues in the mouse PXR ligand binding cavity (Arg-203, Pro-205, Gln-404, Gln-407) to the corresponding residues in the human receptor (Leu-206, Ser-208, His-407, Arg-410, respectively) results in a mouse→human PXR that is activated efficiently by SR12813 but not by PCN, which a rodent-selective PXR ligand (Figure 15). Thus, a "humanized" mouse PXR was created by altering four polar amino acids. This indicates that the relative positions of

critical polar residues are sufficient for achieving selectivity in this receptor. Promiscuity appears to arise from the general hydrophobic nature and relatively small number of polar group within the cavity.

VI.H. Generation of Easily-Solved PXR Crystals

The present invention discloses a substantially pure PXR LBD polypeptide in crystalline form. In a preferred embodiment, exemplified in the Figures and Laboratory Examples, PXR is crystallized with bound ligand. Crystals are formed from PXR LBD polypeptides that are usually expressed by a cell culture, such as *E. coli*. Bromo-, iodo- and substitutions can be included during the preparation of crystal forms and can act as heavy atom substitutions in PXR ligands and in crystals of PXR and the PXR LBD. This method can be advantageous for the phasing of the crystal, which is a crucial, and sometimes limiting, step in solving the three-dimensional structure of a crystallized entity. Thus, the need for generating the heavy metal derivatives traditionally employed in crystallography might be eliminated. After the three-dimensional structure of a PXR or PXR LBD with or without a ligand bound is determined, the resultant three-dimensional structure can be used in computational methods to design synthetic ligands for PXR and other PXR polypeptide fragments. Further activity structure relationships can be determined through routine testing, using assays disclosed herein and known in the art.

VII. Uses of PXR Crystals and the Three-Dimensional Structure of the Ligand Binding Domain of PXR

VII.A. Design and Development of PXR Modulators

The knowledge of the structure of the hPXR ligand binding domain, an aspect of the present invention, provides a tool for investigating the mechanism of action of hPXR and other PXR polypeptides in a subject. For example, various computer models, as described herein, can predict the binding of various substrate molecules to the LBD of hPXR. Upon discovering that such binding in fact takes place, knowledge of the protein structure then

allows design and synthesis of small molecules that mimic the functional binding of the substrate to the LBD of hPXR. This is the method of "rational" drug design, further described herein.

Use of the isolated and purified hPXR crystalline structure of the present invention in rational drug design is thus provided in accordance with the present invention. Additional rational drug design techniques are described in U.S. Patent Nos. 5,834,228 and 5,872,011, incorporated herein in their entirety.

Thus, in addition to the compounds described herein, other sterically similar compounds can be formulated to mimic the key structural regions of PXRs in general, or of hPXR in particular. The generation of a structural functional equivalent can be achieved by the techniques of modeling and chemical design known to those of skill in the art and described herein. It will be understood that all such sterically similar constructs fall within the scope of the present invention.

VII.A.1. Rational Drug Design

The three-dimensional structure of ligand binding hPXR is unprecedented and will greatly aid in the development of new synthetic ligands for a PXR polypeptide, such as PXR agonists and antagonists, including those that bind exclusively to any one of the PXR orthologs. In addition, PXR is well suited to modern methods, including three-dimensional structure elucidation and combinatorial chemistry, such as those disclosed in U.S. Patent No. 5,463,564, incorporated herein by reference. Structure determination using X-ray crystallography is possible because of the solubility properties of the PXR orthologs. Computer programs that use crystallography data when practicing the present invention will enable the rational design of ligands to these receptors. Programs such as RASMOL (Biomolecular Structures Group, Glaxo Wellcome Research & Development Stevenage, Hertfordshire, UK Version 2.6, August 1995, Version 2.6.4, December 1998, Copyright © Roger Sayle 1992-1999) can be used with the atomic structural

coordinates from crystals of the present invention, crystals generated by practicing the invention or crystals used to practice the invention by generating three-dimensional models and/or determining the structures involved in ligand binding. Computer programs such as those sold under the registered trademark INSIGHT II® and such as GRASP (Nicholls et al., (1991) *Proteins* 11: 281-96) allow for further manipulations and the ability to introduce new structures. In addition, high throughput binding and bioactivity assays can be devised using purified recombinant protein and modern reporter gene transcription assays known to those of skill in the art in order to refine the activity of a designed ligand.

A method of identifying modulators of the activity of a PXR polypeptide using rational drug design is thus provided in accordance with the present invention. The method comprises designing a potential modulator for a PXR polypeptide of the present invention that will form non-covalent bonds with amino acids in the ligand binding cavity based upon the crystalline structure of the hPXR LBD polypeptide; synthesizing the modulator; and determining whether the potential modulator modulates the activity of the PXR polypeptide. In a preferred embodiment, the modulator is designed for a hPXR polypeptide. Preferably, the hPXR polypeptide comprises the amino acid sequence of SEQ ID NO: 2, and the hPXR LBD comprises the amino acid sequence SEQ ID NO: 4. The determination of whether the modulator modulates the biological activity of a PXR polypeptide is made in accordance with the screening methods disclosed herein, or by other screening methods known to those of skill in the art. Modulators can be synthesized using techniques known to those of ordinary skill in the art.

In an alternative embodiment, a method of designing a modulator of a PXR polypeptide in accordance with the present invention is disclosed comprising: (a) selecting a candidate PXR ligand; (b) determining which amino acid or amino acids of a PXR polypeptide interact with the ligand using a three-dimensional model of a crystallized hPXR LBD; (c) identifying in a biological assay for PXR activity a degree to which the ligand modulates the activity of the PXR polypeptide; (d) selecting a chemical modification of the

ligand wherein the interaction between the amino acids of the PXR polypeptide and the ligand is predicted to be modulated by the chemical modification; (e) performing the chemical modification on the ligand to form a modified ligand; (f) contacting the modified ligand with the PXR polypeptide; (g) identifying in a biological assay for PXR activity a degree to which the modified ligand modulates the biological activity of the PXR polypeptide ; and (h) comparing the biological activity of the PXR polypeptide in the presence of modified ligand with the biological activity of the PXR polypeptide in the presence of the unmodified ligand, whereby a modulator of a PXR polypeptide is designed.

VII.A.2. Methods for Using the hPXR LBD Structural Coordinates For Molecular Design

For the first time, the present invention permits the use of molecular design techniques to design, select and synthesize chemical entities and compounds, including modulatory compounds, capable of binding to the ligand binding cavity or an accessory binding site of hPXR and the hPXR LBD, in whole or in part. Correspondingly, the present invention also provides for the application of similar techniques in the design of modulators of any PXR polypeptide.

In accordance with a preferred embodiment of the present invention, the structure coordinates of a crystalline hPXR LBD can be used to design compounds that bind to a PXR LBD (more preferably a hPXR LBD) and alter the properties of a PXR LBD (for example, the dimerization or ligand binding ability) in different ways. One aspect of the present invention provides for the design of compounds that act as competitive inhibitors of a PXR polypeptide by binding to all, or a portion of, the binding sites on a PXR LBD. The present invention also provides for the design of compounds that can act as uncompetitive inhibitors of a PXR LBD. These compounds can bind to all, or a portion of, an accessory binding site of a PXR that is already binding its ligand and can, therefore, be more potent and less non-specific than known competitive inhibitors that compete only for the PXR ligand binding cavity.

Similarly, non-competitive inhibitors that bind to and inhibit PXR LBD activity, whether or not it is bound to another chemical entity, can be designed using the PXR LBD structure coordinates of this invention.

A second design approach is to probe a PXR or PXR LBD (preferably a hPXR or hPXR LBD) crystal with molecules comprising a variety of different chemical entities to determine optimal sites for interaction between candidate PXR or PXR LBD modulators and the polypeptide. For example, high resolution X-ray diffraction data collected from crystals saturated with solvent allows the determination of the site where each type of solvent molecule adheres. Small molecules that bind tightly to those sites can then be designed, synthesized and tested for their hPXR modulator activity.

Once a computationally-designed ligand is synthesized using the methods of the present invention or other methods known to those of skill in the art, assays can be used to establish its efficacy of the ligand as a modulator of PXR (preferably hPXR) activity. After such assays, the ligands can be further refined by generating intact PXR, or PXR LBD, crystals with a ligand bound to the PXR. The structure of the ligand can then be further refined using the chemical modification methods described herein and known to those of skill in the art, in order to improve the modulation activity or the binding affinity of the ligand. This process can lead to second generation ligands with improved properties.

Ligands also can be selected that modulate PXR responsive gene transcription by the method of altering the interaction of co-activators and co-repressors with their cognate PXR. For example, agonistic ligands can be selected that block or dissociate a co-repressor from interacting with the PXR, and/or that promote binding or association of a co-activator. Antagonistic ligands can be selected that block co-activator interaction and/or promote co-repressor interaction with a target receptor. Selection can be done via binding assays that screen for designed ligands having the desired modulatory properties. Preferably, interactions of a hPXR polypeptide are targeted. Suitable assays for screening that can be employed, *mutatis mutandis* in the present invention, are described in published PCT international applications

WO 00/037077 and WO 00/025134, incorporated herein by reference in their entirety.

VII.A.3. Methods of Designing PXR LBD Modulator Compounds

The design of candidate substances, also referred to as "compounds" or "candidate compounds", that bind to or inhibit PXR LBD-mediated activity according to the present invention generally involves consideration of two factors. First, the compound must be capable of physically and structurally associating with a PXR LBD. Non-covalent molecular interactions important in the association of a PXR LBD with its substrate include hydrogen bonding, van der Waals interactions and hydrophobic interactions.

Second, the compound must be able to assume a conformation that allows it to associate with a PXR LBD. Although certain portions of the compound will not directly participate in this association with a PXR LBD, those portions can still influence the overall conformation of the molecule. This, in turn, can have a significant impact on potency. Such conformational requirements include the overall three-dimensional structure and orientation of the chemical entity or compound in relation to all or a portion of the binding site, e.g., the ligand binding cavity or an accessory binding site of a PXR LBD, or the spacing between functional groups of a compound comprising several chemical entities that directly interact with a PXR LBD.

The potential modulatory or binding effect of a chemical compound on a PXR LBD can be analyzed prior to its actual synthesis and testing by the use of computer modeling techniques that employ the coordinates of a crystalline hPXR LBD polypeptide of the present invention. If the theoretical structure of the given compound suggests insufficient interaction and association between it and a PXR LBD, synthesis and testing of the compound is obviated. However, if computer modeling indicates a strong interaction, the molecule can then be synthesized and tested for its ability to bind and modulate the activity of a PXR LBD. In this manner, synthesis of unproductive or inoperative compounds can be avoided.

A modulatory or other binding compound of a PXR LBD polypeptide (preferably a hPXR LBD) can be computationally evaluated and designed via a series of steps in which chemical entities or fragments are screened and selected for their ability to associate with the individual binding sites or other areas of a crystalline hPXR LBD polypeptide of the present invention.

One of several methods can be used to screen chemical entities or fragments for their ability to associate with a PXR LBD and, more particularly, with the individual binding sites of a PXR LBD, such as ligand binding cavity or an accessory binding site. This process can begin by visual inspection of, for example, the ligand binding cavity on a computer screen based on the hPXR LBD atomic coordinates in Table 4. Selected fragments or chemical entities can then be positioned in a variety of orientations, or docked, within an individual binding site of a hPXR LBD as defined herein above. Docking can be accomplished using software programs such as those available under the tradenames QUANTA™ (Molecular Simulations Inc., San Diego, California) and SYBYL™ (Tripos, Inc., St. Louis, Missouri), followed by energy minimization and molecular dynamics with standard molecular mechanics forcefields, such as CHARM (Brooks et al., (1983) *J. Comp. Chem.*, 8: 132) and AMBER 5 (Case et al., (1997), AMBER 5, University of California, San Francisco; Pearlman et al., (1995) *Comput. Phys. Commun.* 91: 1-41).

Specialized computer programs can also assist in the process of selecting fragments or chemical entities. These include:

1. GRID™ program, version 17 (Goodford, (1985) *J. Med. Chem.* 28: 849-57), which is available from Molecular Discovery Ltd., Oxford, UK;

2. MCSS™ program (Miranker & Karplus, (1991) *Proteins* 11: 29-34), which is available from Molecular Simulations, Inc., San Diego, California;

3. AUTODOCK™ 3.0 program (Goodsell & Olsen, (1990) *Proteins* 8: 195-202), which is available from the Scripps Research Institute, La Jolla, California;

4. DOCK™ 4.0 program (Kuntz et al., (1992) *J. Mol. Biol.* 161: 269-88), which is available from the University of California, San Francisco, California;

5. FLEX-X™ program (See, Rarey et al., (1996) *J. Comput. Aid. Mol. Des.* 10:41-54), which is available from Tripos, Inc., St. Louis, Missouri;

6. MVP program (Lambert, (1997) in Practical Application of Computer-Aided Drug Design, (Charifson, ed.) Marcel-Dekker, New York, pp. 243-303); and

7. LUDI™ program (Bohm, (1992) *J. Comput. Aid. Mol. Des.*, 6: 61-78), which is available from Molecular Simulations, Inc., San Diego, California.

Once suitable chemical entities or fragments have been selected, they can be assembled into a single compound or modulator. Assembly can proceed by visual inspection of the relationship of the fragments to each other on the three-dimensional image displayed on a computer screen in relation to the structure coordinates of a hPXR LBD. Manual model building using software such as QUANTA™ or SYBYL™ typically follows.

Useful programs to aid one of ordinary skill in the art in connecting the individual chemical entities or fragments include:

1. CAVEAT™ program (Bartlett et al., (1989) *Special Pub., Royal Chem. Soc.* 78: 182-96), which is available from the University of California, Berkeley, California;

2. 3D Database systems, such as MACCS-3D™ system program, which is available from MDL Information Systems, San Leandro, California. This area is reviewed in Martin, (1992) *J. Med. Chem.* 35: 2145-54; and

3. HOOK™ program (Eisen et al., (1994). *Proteins* 19: 199-221), which is available from Molecular Simulations, Inc., San Diego, California.

Instead of proceeding to build a PXR LBD modulator (preferably a hPXR LBD modulator) in a step-wise fashion one fragment or chemical entity at a time as described above, modulatory or other binding compounds can be designed as a whole or *de novo* using the structural coordinates of a crystalline hPXR LBD polypeptide of the present invention and either an empty binding site or optionally including some portion(s) of a known modulator(s). Applicable methods can employ the following software programs:

1. LUDI™ program (Bohm, (1992) *J. Comput. Aid. Mol. Des.*, 6: 61-78), which is available from Molecular Simulations, Inc., San Diego, California;

2. LEGEND™ program (Nishibata & Itai, (1991) *Tetrahedron* 47: 8985); and

3. LEAPFROG™, which is available from Tripos Associates, St. Louis, Missouri.

Other molecular modeling techniques can also be employed in accordance with this invention. See, e.g., Cohen et al., (1990) *J. Med. Chem.* 33: 883-94. See also, Navia & Murcko, (1992) *Curr. Opin. Struc. Biol.* 2: 202-10; U.S. Patent No. 6,008,033, herein incorporated by reference.

Once a compound has been designed or selected by the above methods, the efficiency with which that compound can bind to a PXR LBD can be tested and optimized by computational evaluation. By way of particular example, a compound that has been designed or selected to function as a hPXR LBD modulator should also preferably traverse a volume not overlapping that occupied by the binding site when it is bound to its native ligand. Additionally, an effective PXR LBD modulator should preferably demonstrate a relatively small difference in energy between its bound and free states (i.e., a small deformation energy of binding). Thus, the most efficient PXR LBD modulators should preferably be designed with a deformation energy of binding of not greater than about 10 kcal/mole, and preferably, not greater than 7 kcal/mole. It is possible for PXR LBD modulators to interact with the polypeptide in more than one conformation that is similar in overall binding energy. In those cases, the deformation energy of binding is taken to be the difference between the energy of the free compound and the average energy of the conformations observed when the modulator binds to the polypeptide.

A compound designed or selected as binding to a PXR polypeptide (preferably a hPXR LBD polypeptide) can be further computationally optimized so that in its bound state it would preferably lack repulsive electrostatic interaction with the target polypeptide. Such non-complementary (e.g., electrostatic) interactions include repulsive charge-charge, dipole-dipole

and charge-dipole interactions. Specifically, the sum of all electrostatic interactions between the modulator and the polypeptide when the modulator is bound to a PXR LBD preferably make a neutral or favorable contribution to the enthalpy of binding.

Specific computer software is available in the art to evaluate compound deformation energy and electrostatic interaction. Examples of programs designed for such uses include:

1. Gaussian 98™, which is available from Gaussian, Inc., Pittsburgh, Pennsylvania;

2. AMBER™ program, version 6.0, which is available from the University of California at San Francisco;

3. QUANTA™ program, which is available from Molecular Simulations, Inc., San Diego, California;

4. CHARMm® program, which is available from Molecular Simulations, Inc., San Diego, California; and

4. INSIGHT II® program, which is available from Molecular Simulations, Inc., San Diego, California.

These programs can be implemented using a suitable computer system. Other hardware systems and software packages will be apparent to those skilled in the art after review of the disclosure of the present invention presented herein.

Once a PXR LBD modulating compound has been optimally selected or designed, as described above, substitutions can then be made in some of its atoms or side groups in order to improve or modify its binding properties. Generally, initial substitutions are conservative, i.e., the replacement group will have approximately the same size, shape, hydrophobicity and charge as the original group. It should, of course, be understood that components known in the art to alter conformation should be avoided. Such substituted chemical compounds can then be analyzed for efficiency of fit to a PXR LBD binding site using the same computer-based approaches described in detail above.

VII.B. Distinguishing Between PXR Isoforms and Orthologs

The present invention discloses the ability to generate new synthetic ligands to distinguish between PXR isoforms and orthologs. As described herein, computer-designed ligands can be generated that distinguish between binding isoforms and orthologs, thereby allowing the generation of species specific, tissue specific or function specific ligands. The atomic structural coordinates disclosed in the present invention reveal structural details unique to hPXR. These structural details can be exploited when a novel ligand is designed using the methods of the present invention or other ligand design methods known in the art. The structural features that differentiate a hPXR from a mouse PXR and one isoform from another can be targeted in ligand design. Thus, for example, a ligand can be designed that will recognize a particular PXR isoform or ortholog, while not interacting with other PXR isoforms or orthologs, or even with moieties having similar structural features. Prior to the disclosure of the present invention, a detailed understanding of the differences between PXR orthoforms and the ability to target a particular PXR isoform or ortholog was unattainable.

VII.C. Method of Screening for Chemical and Biological Modulators of the Biological Activity of hPXR

A candidate substance identified according to a screening assay of the present invention has an ability to modulate the biological activity of a PXR polypeptide or a PXR LBD polypeptide. In a preferred embodiment, such a candidate compound can have utility in the treatment of disorders and conditions associated with the biological activity of a hPXR or a hPXR LBD polypeptide, including, but not limited to, hPXR and hPXR LBD-based drug-drug interactions, hPXR and hPXR LBD-based drug resistance, individualized treatment of disease due to polymorphisms, liver cholestasis and other degenerative liver disorders.

In a cell-free system, the method comprises the steps of establishing a control system comprising a hPXR polypeptide and a ligand which is capable of binding to the polypeptide; establishing a test system comprising a hPXR

polypeptide, the ligand, and a candidate compound; and determining whether the candidate compound modulates the activity of the polypeptide by comparison of the test and control systems. A representative ligand comprises a fatty acid or other small molecule, and in this embodiment, the biological activity or property screened includes binding affinity.

In another embodiment of the invention, a form of a hPXR polypeptide or a catalytic or immunogenic fragment or oligopeptide thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such a screening can be affixed to a solid support. The formation of binding complexes, between a hPXR polypeptide and the agent being tested, will be detected. In a preferred embodiment, the hPXR polypeptide has an amino acid sequence of SEQ ID NO: 2. When a hPXR LBD polypeptide is employed, a preferred embodiment will include a hPXR polypeptide having the amino acid sequence of SEQ ID NO: 4.

Another technique for drug screening which can be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO 84/03564, herein incorporated by reference. In this method, as applied to a polypeptide of the present invention, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with the polypeptide, or fragments thereof. Bound polypeptide is then detected by methods well known to those of skill in the art. The polypeptide can also be placed directly onto plates for use in the aforementioned drug screening techniques.

In yet another embodiment, a method of screening for a modulator of a hPXR polypeptide or a hPXR LBD polypeptide comprises: providing a library of test samples; contacting a hPXR polypeptide or a hPXR LBD polypeptide with each test sample; detecting an interaction between a test sample and a hPXR polypeptide or a hPXR LBD polypeptide; identifying a test sample that interacts with a hPXR polypeptide or a hPXR LBD polypeptide; and isolating a

test sample that interacts with a hPXR polypeptide or a hPXR LBD polypeptide.

In each of the foregoing embodiments, an interaction can be detected spectrophotometrically, radiologically or immunologically. An interaction between a hPXR polypeptide or a hPXR LBD polypeptide and a test sample can also be quantified using methodology known to those of skill in the art. In another embodiment, the hPXR polypeptide and the hPXR LBD is in crystalline form.

In accordance with the present invention there is also provided a rapid and high throughput screening method that relies on the methods described above. This screening method comprises separately contacting each of a plurality of substantially identical samples with a hPXR polypeptide or a hPXR LBD and detecting a resulting binding complex. In such a screening method the plurality of samples preferably comprises more than about 10^4 samples, or more preferably comprises more than about 5×10^4 samples.

VII.D. Method of Identifying Compounds Which Inhibit Ligand Binding

Using the crystal structures and ligand orientations, disclosed for the first time in the present invention, it is possible to design test compounds that inhibit binding of ligands normally bound by a PXR polypeptide.

In one aspect of the present invention, an assay method for identifying a compound that inhibits binding of a ligand to a PXR polypeptide is disclosed. A known ligand of hPXR can be used in the assay method as the ligand against which the inhibition by a test compound is gauged. SR12813 is a preferred ligand in the assay method. The method comprises (a) incubating a PXR polypeptide with a ligand in the presence of a test inhibitor compound; (b) determining an amount of ligand that is bound to the PXR polypeptide, wherein decreased binding of ligand to the PXR polypeptide in the presence of the test inhibitor compound relative to binding in the absence of the test inhibitor compound is indicative of inhibition; and (c) identifying the test compound as an inhibitor of ligand binding if decreased ligand binding is observed. Preferably, the ligand is SR12813.

In another aspect of the present invention, the disclosed assay method can be used in the structural refinement of candidate PXR inhibitors. For example, multiple rounds of optimization can be followed by gradual structural changes in a strategy of inhibitor design. A strategy such as this is made possible by the disclosure of the coordinates of the hPXR LBD and the disclosure of the orientation of a ligand of PXR, SR12813.

VII.E. Design of PXR Isoform and Ortholog Modulators

The hPXR crystal structure of the present invention can be used to generate modulators of other PXR isoforms or orthologs, such as mouse PXR (Swiss-Prot Accession No. O54915). Analysis of the disclosed crystal structure can provide a guide for designing modulators of PXR isoforms or orthologs. Purely for purposes of explanation, the development of a mouse PXR modulator will be considered herein below. It will be apparent to those of skill in the art, and explicitly noted here, that the following discussion will be applicable *mutatis mutandis* to PXR isoforms and other PXR orthologs, including rat PXR (Swiss-Prot Accession No. Q9R1A7).

Absent the crystal structure of the present invention, researchers would be required to design mouse PXR modulators *de novo*. The present invention, however, addresses this problem by providing insights into the binding cavity of hPXR, which can be extended, due to significant structural similarity with other PXR isoforms and orthologs, to the binding cavity of, for example, mouse PXR. An evaluation of the binding cavity of hPXR indicates that a potential mouse PXR modulator would meet a broad set of general criteria. Broadly, it can be stated that, based on the crystal structure of hPXR, a potent mouse PXR ligand would require several general features including: (a) a hydrophobic binding cavity; and (b) the ability to adopt a conformation that is complementary to the shape of the binding cavity.

Using the discerned structural similarities and differences between PXR isoforms and orthologs, as represented and predicted based on the crystal structure of the present invention and homology models, a mouse PXR modulator can be designed. For example, based on an evaluation of a

homology model of mouse PXR, which is derived from the hPXR crystal structure, it is expected that a potent ligand would need similar characteristics as listed above for a compound recognized by hPXR. Additional modifications can be included, based on the disclosed structure, which are predicted to further define a modulator specific for mouse PXR over other orthologs. Thus, the disclosed crystal structure of hPXR can be useful when designing modulators of mouse PXR and other orthologs and isoforms.

An additional aspect of the present invention, then, is to provide a technique for predicting the differences in the metabolism of a drug between mice (or other species of interest) and humans. As described further herein, the "humanization" of the mouse PXR via four point mutations demonstrate that the structures disclosed herein can be employed as a predictor of mouse PXR function. A mouse PXR model, based on the structure of human PXR, as disclosed herein and co-crystallized and refined with known pharmacophores, can provide needed insight into PXR activity and mechanisms. Notably, a "humanized" mouse PXR can provide valuable insight into the metabolism of drugs in humans. Particularly, the ability to generate such a system offers the potential to generate PXR drug metabolism data in a mouse system that is more predictive for human systems, than PXR metabolic data generated based only on a mouse model. A "humanized" mouse PXR also facilitates insights into existent drug metabolism, and the development of human drugs and therapeutics with greater efficacies, using a convenient mouse model system.

VIII. Design, Preparation and Structural Analysis of hPXR and hPXR LBD Mutants and Structural Equivalents

The present invention provides for the generation of PXR and PXR mutants (preferably hPXR and hPXR LBD mutants), and the ability to solve the crystal structures of those that crystallize. More particularly, through the provision of the three-dimensional structure of a hPXR LBD, desirable sites for mutation can be identified, based on analysis of the three-dimensional hPXR LBD structure coordinates provided herein.

The structure coordinates of a hPXR LBD provided in accordance with the present invention also facilitate the identification of related proteins or enzymes analogous to hPXR in function, structure or both, (for example, a mouse PXR), which can lead to novel therapeutic modes for treating or preventing a range of disease states.

VIII.A. Sterically Similar Compounds

A further aspect of the present invention is that sterically similar compounds can be formulated to mimic the key portions of a PXR LBD structure. Such compounds are functional equivalents. The generation of a structural furictional equivalent can be achieved by the techniques of modeling and chemical design known to those of skill in the art and described herein. Modeling and chemical design of PXR and PXR LBD structural equivalents can be based on the structure coordinates of a crystalline hPXR LBD polypeptide of the present invention. It will be understood that all such sterically similar constructs fall within the scope of the present invention.

VIII.B. PXR Polypeptides

The generation of chimeric PXR polypeptides is also an aspect of the present invention. Such a chimeric polypeptide can comprise a PXR LBD polypeptide or a portion of a PXR LBD, (e.g. a hPXR LBD) which is fused to a candidate polypeptide or a suitable region of the candidate polypeptide, for example a PXR expressed in mouse or other species. Throughout the present disclosure it is intended that the term "mutant" encompass not only mutants of a PXR LBD polypeptide but chimeric proteins generated using a PXR LBD as well. It is thus intended that the following discussion of mutant PXR LBDs apply *mutatis mutandis* to chimeric PXR and PXR LBD polypeptides and to structural equivalents thereof.

In accordance with the present invention, a mutation can be directed to a particular site or combination of sites of a wild-type PXR LBD. For example, an accessory binding site or the binding cavity can be chosen for mutagenesis. Similarly, a residue having a location on, at or near the surface

of the polypeptide can be replaced, resulting in an altered surface charge of one or more charge units, as compared to the wild-type PXR and PXR LBD. Alternatively, an amino acid residue in a PXR or a PXR LBD can be chosen for replacement based on its hydrophilic or hydrophobic characteristics.

Such mutants can be characterized by any one of several different properties as compared with the wild-type PXR LBD. For example, such mutants can have an altered surface charge of one or more charge units, or can have an increase in overall stability. Other mutants can have altered substrate specificity in comparison with, or a higher specific activity than, a wild-type PXR or PXR LBD.

PXR and PXR LBD mutants of the present invention can be generated in a number of ways. For example, the wild-type sequence of a PXR or a PXR LBD can be mutated at those sites identified using this invention as desirable for mutation, by means of oligonucleotide-directed mutagenesis or other conventional methods, such as deletion. Alternatively, mutants of a PXR or a PXR LBD can be generated by the site-specific replacement of a particular amino acid with an unnaturally occurring amino acid. In addition, PXR or PXR LBD mutants can be generated through replacement of an amino acid residue, for example, a particular cysteine or methionine residue, with selenocysteine or selenomethionine. This can be achieved by growing a host organism capable of expressing either the wild-type or mutant polypeptide on a growth medium depleted of either natural cysteine or methionine (or both) but enriched in selenocysteine or selenomethionine (or both).

Mutations can be introduced into a DNA sequence coding for a PXR or a PXR LBD using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites. Mutations can be generated in the full-length DNA sequence of a PXR or a PXR LBD or in any sequence coding for polypeptide fragments of a PXR or a PXR LBD.

According to the present invention, a mutated PXR or PXR LBD DNA sequence produced by the methods described above, or any alternative methods known in the art, can be expressed using an expression vector. An

expression vector, as is well known to those of skill in the art, typically includes elements that permit autonomous replication in a host cell independent of the host genome, and one or more phenotypic markers for selection purposes. Either prior to or after insertion of the DNA sequences surrounding the desired PXR or PXR LBD mutant coding sequence, an expression vector also will include control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and, optionally, a repressor gene or various activator genes and a signal for termination. In some embodiments, where secretion of the produced mutant is desired, nucleotides encoding a "signal sequence" can be inserted prior to a PXR or a PXR LBD mutant coding sequence. For expression under the direction of the control sequences, a desired DNA sequence must be operatively linked to the control sequences; that is, the sequence must have an appropriate start signal in front of the DNA sequence encoding the PXR or PXR LBD mutant, and the correct reading frame to permit expression of that sequence under the control of the control sequences and production of the desired product encoded by that PXR or PXR LBD sequence must be maintained.

Any of a wide variety of well-known available expression vectors can be useful to express a mutated PXR or PXR LBD coding sequences of this invention and generated as described in Laboratory Example 3. These expression vectors can be used in the techniques disclosed in Laboratory Examples 1 and 3 and can include, for example, vectors comprising segments of chromosomal, non-chromosomal and synthetic DNA sequences, such as various known derivatives of SV40, known bacterial plasmids, e.g., plasmids from *E. coli* including col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, e.g., RP4, phage DNAs, e.g., the numerous derivatives of phage λ, e.g., NM 989, and other DNA phages, e.g., M13 and filamentous single stranded DNA phages, yeast plasmids and vectors derived from combinations of plasmids and phage DNAs, such as plasmids which have been modified to employ phage DNA or other expression control sequences. In a preferred embodiment of this invention, the *E. coli* vector pRSETA, including a T7-based expression system, is employed.

In addition, any of a wide variety of expression control sequences—sequences that control the expression of a DNA sequence when operatively linked to it—can be used in these vectors to express the mutated DNA sequences according to this invention. Such useful expression control sequences, include, for example, the early and late promoters of SV40 for animal cells, the lac system, the trp system the TAC or TRC system, the major operator and promoter regions of phage λ, the control regions of fd coat protein, all for *E. coli*, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors for yeast, and other sequences known to control the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof.

A wide variety of hosts are also useful for producing mutated hPXR and hPXR LBD polypeptides according to this invention. These hosts include, for example, bacteria, such as *E. coli*, *Bacillus* and *Streptomyces*, fungi, such as yeasts, and animal cells, such as CHO and COS-1 cells, plant cells, insect cells, such as Sf9 cells, and transgenic host cells.

It should be understood that not all expression vectors and expression systems function in the same way to express mutated DNA sequences of this invention, and to produce modified PXR and PXR LBD polypeptides or PXR or PXR LBD mutants. Neither do all hosts function equally well with the same expression system. One of skill in the art can, however, make a selection among these vectors, expression control sequences and hosts without undue experimentation and without departing from the scope of this invention. For example, an important consideration in selecting a vector will be the ability of the vector to replicate in a given host. The copy number of the vector, the ability to control that copy number, and the expression of any other proteins encoded by the vector, such as antibiotic markers, should also be considered.

In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the system, its controllability and its compatibility with the DNA sequence encoding a modified PXR or PXR LBD polypeptide of this invention, with

particular regard to the formation of potential secondary and tertiary structures.

Hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of a modified PXR or PXR LBD to them, their ability to express mature products, their ability to fold proteins correctly, their fermentation requirements, the ease of purification of a modified PXR or PXR LBD and safety. Within these parameters, one of skill in the art can select various vector/expression control system/host combinations that will produce useful amounts of a mutant PXR or PXR LBD. A mutant PXR or PXR LBD produced in these systems can be purified by a variety of conventional steps and strategies, including those used to purify the wild-type PXR or PXR LBD.

Once a PXR LBD mutation(s) has been generated in the desired location, such as an ligand binding or dimerization site, the mutants can be tested for any one of several properties of interest. For example, mutants can be screened for an altered charge at physiological pH. This is determined by measuring the mutant PXR or PXR LBD isoelectric point (pI) and comparing the observed value with that of the wild-type parent. Isoelectric point can be measured by gel-electrophoresis according to the method of Wellner (Wellner, (1971) *Anal. Chem.* 43: 597). A mutant PXR or PXR LBD polypeptide containing a replacement amino acid located at the surface of the enzyme, as provided by the structural information of this invention, can lead to an altered surface charge and an altered pI.

VIII.C. Generation of an Engineered PXR or PXR LBD Mutant

In another aspect of the present invention, a unique PXR or PXR LBD polypeptide can be generated. Such a mutant can facilitate purification and can facilitate the study of the ligand binding abilities of a PXR polypeptide.

As used in the following discussion, the terms "engineered PXR", "engineered PXR LDB", "PXR mutant", and "PXR LBD mutant" refers to polypeptides having amino acid sequences which contain at least one mutation in the wild-type sequence. The terms also refer to PXR and PXR LBD polypeptides which are capable of exerting a biological effect in that they

comprise all or a part of the amino acid sequence of an engineered PXR or PXR LBD mutant polypeptide of the present invention, or cross-react with antibodies raised against an engineered PXR or PXR LBD mutant polypeptide, or retain all or some or an enhanced degree of the biological activity of the engineered PXR or PXR LBD mutant amino acid sequence or protein. Such biological activity can include lipid binding in general, and fatty acid binding in particular.

The terms "engineered PXR LBD" and "PXR LBD mutant" also includes analogs of an engineered PXR LBD or PXR LBD mutant polypeptide. By "analog" is intended that a DNA or polypeptide sequence can contain alterations relative to the sequences disclosed herein, yet retain all or some or an enhanced degree of the biological activity of those sequences. Analogs can be derived from genomic nucleotide sequences or from other organisms, or can be created synthetically. Those of skill in the art will appreciate that other analogs, as yet undisclosed or undiscovered, can be used to design and/or construct PXR LBD or PXR LBD mutant analogs. There is no need for an engineered PXR LBD or PXR LBD mutant polypeptide to comprise all or substantially all of the amino acid sequence of SEQ ID NOs: 2 or 4. Shorter or longer sequences are anticipated to be of use in the invention; shorter sequences are herein referred to as "segments". Thus, the terms "engineered PXR LBD" and "PXR LBD mutant" also includes fusion, chimeric or recombinant engineered PXR LBD or PXR LBD mutant polypeptides and proteins comprising sequences of the present invention. Methods of preparing such proteins are disclosed herein above and are known in the art.

VIII.D. Sequence Similarity and Identity

As used herein, the term "substantially similar" means that a particular sequence varies from nucleic acid sequence of SEQ ID NOs: 1 or 3, or the amino acid sequence of SEQ ID NOs: 2 or 4 by one or more deletions, substitutions, or additions, the net effect of which is to retain at least some of biological activity of the natural gene, gene product, or sequence. Such sequences include "mutant" or "polymorphic" sequences, or sequences in

which the biological activity and/or the physical properties are altered to some degree but retains at least some or an enhanced degree of the original biological activity and/or physical properties. In determining nucleic acid sequences, all subject nucleic acid sequences capable of encoding substantially similar amino acid sequences are considered to be substantially similar to a reference nucleic acid sequence, regardless of differences in codon sequences or substitution of equivalent amino acids to create biologically functional equivalents.

VIII.D.1. Sequences That are Substantially Identical to an Engineered PXR or PXR LBD Mutant Sequence of the Present Invention

Nucleic acids that are substantially identical to a nucleic acid sequence of an engineered PXR or PXR LBD mutant of the present invention, e.g. allelic variants, genetically altered versions of the gene, etc., bind to an engineered PXR or PXR LBD mutant sequence under stringent hybridization conditions. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes. The source of homologous genes can be any species, e.g. primate species; rodents, such as rats and mice, canines, felines, bovines, equines, yeast, nematodes, etc.

Between mammalian species, e.g. human and mouse, homologs have substantial sequence similarity, i.e. at least 75% sequence identity between nucleotide sequences. Sequence similarity is calculated based on a reference sequence, which can be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, etc. A reference sequence will usually be at least about 18 nucleotides(nt) long, more usually at least about 30 nt long, and can extend to the complete sequence that is being compared. Algorithms for sequence analysis are known in the art, such as BLAST, described in Altschul et al., (1990) *J. Mol. Biol.* 215: 403-10.

Percent identity or percent similarity of a DNA or peptide sequence can be determined, for example, by comparing sequence information using the GAP computer program, available from the University of Wisconsin Geneticist

Computer Group. The GAP program utilizes the alignment method of Needleman et al., (1970) *J. Mol. Biol.* 48: 443, as revised by Smith et al., (1981) *Adv. Appl. Math.* 2:482. Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) that are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred parameters for the GAP program are the default parameters, which do not impose a penalty for end gaps. See, e.g., Schwartz et al., eds., (1979), Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, pp. 357-358, and Gribskov et al., (1986) *Nucl. Acids. Res.* 14: 6745.

The term "similarity" is contrasted with the term "identity". Similarity is defined as above; "identity", however, means a nucleic acid or amino acid sequence having the same amino acid at the same relative position in a given family member of a gene family. Homology and similarity are generally viewed as broader terms than the term identity. Biochemically similar amino acids, for example leucine/isoleucine or glutamate/aspartate, can be present at the same position—these are not identical per se, but are biochemically "similar." As disclosed herein, these are referred to as conservative differences or conservative substitutions. This differs from a conservative mutation at the DNA level, which changes the nucleotide sequence without making a change in the encoded amino acid, e.g. TCC to TCA, both of which encode serine.

As used herein, DNA analog sequences are "substantially identical" to specific DNA sequences disclosed herein if: (a) the DNA analog sequence is derived from coding regions of the nucleic acid sequence shown in SEQ ID NOs: 1 or 3; or (b) the DNA analog sequence is capable of hybridization with DNA sequences of (a) under stringent conditions and which encode a biologically active hPXR or hPXR LBD gene product; or (c) the DNA sequences are degenerate as a result of alternative genetic code to the DNA analog sequences defined in (a) and/or (b). Substantially identical analog proteins and nucleic acids will have between about 70% and 80%, preferably between about 81% to about 90% or even more preferably between about

91% and 99% sequence identity with the corresponding sequence of the native protein or nucleic acid. Sequences having lesser degrees of identity but comparable biological activity are considered to be equivalents.

As used herein, "stringent conditions" means conditions of high stringency, for example 6X SSC, 0.2% polyvinylpyrrolidone, 0.2% Ficoll, 0.2% bovine serum albumin, 0.1% sodium dodecyl sulfate, 100 μg/ml salmon sperm DNA and 15% formamide at 68°C. For the purposes of specifying additional conditions of high stringency, preferred conditions are salt concentration of about 200 mM and temperature of about 45°C. One example of such stringent conditions is hybridization at 4X SSC, at 65°C, followed by a washing in 0.1X SSC at 65°C for one hour. Another exemplary stringent hybridization scheme uses 50% formamide, 4X SSC at 42°C.

In contrast, nucleic acids having sequence similarity are detected by hybridization under lower stringency conditions. Thus, sequence identity can be determined by hybridization under lower stringency conditions, for example, at 50°C or higher and 0.1X SSC (9 mM NaCl/0.9 mM sodium citrate) and the sequences will remain bound when subjected to washing at 55°C in 1X SSC.

VIII.D.2. Complementarity and Hybridization to an Engineered PXR or PXR LBD Mutant Sequence

As used herein, the term "complementary sequences" means nucleic acid sequences which are base-paired according to the standard Watson-Crick complementarity rules. The present invention also encompasses the use of nucleotide segments that are complementary to the sequences of the present invention.

Hybridization can also be used for assessing complementary sequences and/or isolating complementary nucleotide sequences. As discussed above, nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be

readily appreciated by those skilled in the art. Stringent temperature conditions will generally include temperatures in excess of about 30°C, typically in excess of about 37°C, and preferably in excess of about 45°C. Stringent salt conditions will ordinarily be less than about 1,000 mM, typically less than about 500 mM, and preferably less than about 200 mM. However, the combination of parameters is much more important than the measure of any single parameter. See, e.g., Wetmur & Davidson, (1968) *J. Mol. Biol.* 31: 349-70. Determining appropriate hybridization conditions to identify and/or isolate sequences containing high levels of homology is well known in the art. See, e.g., Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York.

VIII.D.3. Functional Equivalents of an Engineered PXR or PXR LBD Mutant Nucleic Acid Sequence of the Present Invention

As used herein, the term "functionally equivalent codon" is used to refer to codons that encode the same amino acid, such as the ACG and AGU codons for serine. Human PXR or hPXR LBD-encoding nucleic acid sequences comprising SEQ ID NOs: 1 and 3, which have functionally equivalent codons, are covered by the present invention. Thus, when referring to the sequence example presented in SEQ ID NOs: 1 and 3, applicants contemplate substitution of functionally equivalent codons into the sequence example of SEQ ID NOs: 1 and 3. Thus, applicants are in possession of amino acid and nucleic acids sequences which include such substitutions but which are not set forth herein in their entirety for convenience.

It will also be understood by those of skill in the art that amino acid and nucleic acid sequences can include additional residues, such as additional N- or C-terminal amino acids or 5' or 3' nucleic acid sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence retains biological protein activity where polypeptide expression is concerned. The addition of terminal sequences particularly applies to

nucleic acid sequences which can, for example, include various non-coding sequences flanking either of the 5' or 3' portions of the coding region or can include various internal sequences, i.e., introns, which are known to occur within genes.

VIII.D.4. Biological Equivalents

The present invention envisions and includes biological equivalents of an engineered PXR or PXR LBD mutant polypeptide of the present invention. The term "biological equivalent" refers to proteins having amino acid sequences which are substantially identical to the amino acid sequence of an engineered PXR LBD mutant of the present invention and which are capable of exerting a biological effect in that they are capable of binding DNA moieties or cross-reacting with anti-PXR or PXR LBD mutant antibodies raised against an engineered mutant PXR or PXR LBD polypeptide of the present invention.

For example, certain amino acids can be substituted for other amino acids in a protein structure without appreciable loss of interactive capacity with, for example, structures in the nucleus of a cell. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence (or the nucleic acid sequence encoding it) to obtain a protein with the same, enhanced, or antagonistic properties. Such properties can be achieved by interaction with the normal targets of the protein, but this need not be the case, and the biological activity of the invention is not limited to a particular mechanism of action. It is thus in accordance with the present invention that various changes can be made in the amino acid sequence of an engineered PXR or PXR LBD mutant polypeptide of the present invention or its underlying nucleic acid sequence without appreciable loss of biological utility or activity.

Biologically equivalent polypeptides, as used herein, are polypeptides in which certain, but not most or all, of the amino acids can be substituted. Thus, when referring to the sequence examples presented in SEQ ID NOs: 1 and 3, applicants envision substitution of codons that encode biologically

equivalent amino acids, as described herein, into the sequence example of SEQ ID NOs: 1 and 3, respectively. Thus, applicants are in possession of amino acid and nucleic acids sequences which include such substitutions but which are not set forth herein in their entirety for convenience.

Alternatively, functionally equivalent proteins or peptides can be created via the application of recombinant DNA technology, in which changes in the protein structure can be engineered, based on considerations of the properties of the amino acids being exchanged, e.g. substitution of Ile for Leu. Changes designed by man can be introduced through the application of site-directed mutagenesis techniques, e.g., to introduce improvements to the antigenicity of the protein or to test an engineered PXR or PXR LBD mutant polypeptide of the present invention in order to modulate DNA-binding, lipid-binding or other activity, at the molecular level.

Amino acid substitutions, such as those which might be employed in modifying an engineered PXR or PXR LBD mutant polypeptide of the present invention are generally, but not necessarily, based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. An analysis of the size, shape and type of the amino acid side-chain substituents reveals that arginine, lysine and histidine are all positively charged residues; that alanine, glycine and serine are all of similar size; and that phenylalanine, tryptophan and tyrosine all have a generally similar shape. Therefore, based upon these considerations, arginine, lysine and histidine; alanine, glycine and serine; and phenylalanine, tryptophan and tyrosine; are defined herein as biologically functional equivalents. Other biologically functionally equivalent changes will be appreciated by those of skill in the art. It is implicit in the above discussion, however, that one of skill in the art can appreciate that a radical, rather than a conservative substitution is warranted in a given situation. Non-conservative substitutions in engineered mutant PXR or PXR LBD polypeptides of the present invention are also an aspect of the present invention.

In making biologically functional equivalent amino acid substitutions, the hydropathic index of amino acids can be considered. Each amino acid

has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics, these are: isoleucine (+ 4.5); valine (+ 4.2); leucine (+ 3.8); phenylalanine (+ 2.8); cysteine (+ 2.5); methionine (+ 1.9); alanine (+ 1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte & Doolittle, (1982), *J. Mol. Biol.* 157: 105-132, incorporated herein by reference). It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ±2 of the original value is preferred, those which are within ±1 of the original value are particularly preferred, and those within ±0.5 of the original value are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent No. 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, i.e. with a biological property of the protein. It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent protein.

As detailed in U.S. Patent No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+ 3.0); lysine (+ 3.0); aspartate (+ 3.0±1); glutamate (+ 3.0±1); serine (+ 0.3); asparagine (+ 0.2); glutamine (+ 0.2); glycine (0); threonine (-0.4); proline (-0.5±1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).

In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ±2 of the

original value is preferred, those which are within ±1 of the original value are particularly preferred, and those within ±0.5 of the original value are even more particularly preferred.

While discussion has focused on functionally equivalent polypeptides arising from amino acid changes, it will be appreciated that these changes can be effected by alteration of the encoding DNA, taking into consideration also that the genetic code is degenerate and that two or more codons can code for the same amino acid.

Thus, it will also be understood that this invention is not limited to the particular amino acid and nucleic acid sequences of SEQ ID NOs: 1-4. Recombinant vectors and isolated DNA segments can therefore variously include an engineered hPXR or hPXR LBD mutant polypeptide-encoding region itself, include coding regions bearing selected alterations or modifications in the basic coding region, or include larger polypeptides which nevertheless comprise a hPXR or hPXR LBD mutant polypeptide-encoding regions or can encode biologically functional equivalent proteins or polypeptides which have variant amino acid sequences. Biological activity of an engineered hPXR or hPXR LBD mutant polypeptide can be determined, for example, by lipid-binding assays known to those of skill in the art.

The nucleic acid segments of the present invention, regardless of the length of the coding sequence itself, can be combined with other DNA sequences, such as promoters, enhancers, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length can vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length can be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, nucleic acid fragments can be prepared which include a short stretch complementary to a nucleic acid sequence set forth in SEQ ID NOs: 1 and 3, such as about 10 nucleotides, and which are up to 10,000 or 5,000 base pairs in length. DNA segments with total lengths of about 4,000, 3,000, 2,000, 1,000, 500, 200, 100, and about 50 base pairs in length are also useful.

The DNA segments of the present invention encompass biologically functional equivalents of engineered PXR or PXR LBD mutant polypeptides. Such sequences can rise as a consequence of codon redundancy and functional equivalency that are known to occur naturally within nucleic acid sequences and the proteins thus encoded. Alternatively, functionally equivalent proteins or polypeptides can be created via the application of recombinant DNA technology, in which changes in the protein structure can be engineered, based on considerations of the properties of the amino acids being exchanged. Changes can be introduced through the application of site-directed mutagenesis techniques, e.g., to introduce improvements to the antigenicity of the protein or to test variants of an engineered PXR or PXR LBD mutant of the present invention in order to examine the degree of lipid-binding activity, or other activity at the molecular level. Various site-directed mutagenesis techniques are known to those of skill in the art and can be employed in the present invention.

The invention further encompasses fusion proteins and peptides wherein an engineered PXR or PXR LBD mutant coding region of the present invention is aligned within the same expression unit with other proteins or peptides having desired functions, such as for purification or immunodetection purposes.

Recombinant vectors form important further aspects of the present invention. Particularly useful vectors are those in which the coding portion of the DNA segment is positioned under the control of a promoter. The promoter can be that naturally associated with a PXR gene, as can be obtained by isolating the 5' non-coding sequences located upstream of the coding segment or exon, for example, using recombinant cloning and/or PCR technology and/or other methods known in the art, in conjunction with the compositions disclosed herein.

In other embodiments, certain advantages will be gained by positioning the coding DNA segment under the control of a recombinant, or heterologous, promoter. As used herein, a recombinant or heterologous promoter is a promoter that is not normally associated with a PXR gene in its natural

environment. Such promoters can include promoters isolated from bacterial, viral, eukaryotic, or mammalian cells. Naturally, it will be important to employ a promoter that effectively directs the expression of the DNA segment in the cell type chosen for expression. The use of promoter and cell type combinations for protein expression is generally known to those of skill in the art of molecular biology (See, e.g., Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York, specifically incorporated herein by reference). The promoters employed can be constitutive or inducible and can be used under the appropriate conditions to direct high level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant proteins or peptides. One preferred promoter system contemplated for use in high-level expression is a T7 promoter-based system.

IX. Uses of PXR Mutants

The PXR mutants disclosed herein have a variety of applications, including in the screening of components for PXR/SXR activation using the cell-free reporter gene assay methods disclosed herein above, and using whole animal models. The PXR mutants can also be used in cell-free, cell-based and whole animal assay methods for bioavailability of compounds and for toxicology analysis. Additionally, PXR mutants can be employed in crystallizations, screening for changes in ligand activation, screening for species-specific changes in ligand activation and screening for changes in oligomerization state both with and without ligand.

IX.A. Evaluating a Candidate Therapeutic Agent in Humans Using a Mouse Model System

A mutant PXR comprising a "humanized" mouse can find particular utility as a model for the study of drug metabolism in humans. Such a model can be more informative regarding drug metabolism in humans than studies using unmodified mouse PXR as a predictive model for human drug

metabolism. It also offers the convenience and advantages of employing a mouse model system to study a human protein.

In a preferred embodiment, a method of evaluating a candidate therapeutic agent in humans using a mouse model system comprises the following steps. Atomic coordinates of a human PXR ligand binding domain are provided to a computerized modeling system. Preferred coordinates are supplied in Table 4 of the present disclosure. Based on the human PXR atomic coordinates, a candidate therapeutic agent that fits spatially into a binding cavity or on the surface of a human PXR ligand binding domain is modeled. Modeling of a candidate therapeutic agent can conveniently employ a computer-based approach. Computer programs, such as the INSIGHT II® program (Molecular Simulations, Inc., San Diego, California) can be employed to create three-dimensional structures of candidate agents. After a candidate therapeutic agent structure is generated, a computer program (which can comprise the appropriate design and structure evaluation modules), can be employed to preliminarily evaluate the candidate agent for use as a therapeutic. The candidate agent can be docked with a site on the PXR, for example, and the efficiency of the interaction can be computationally evaluated for various physical properties, such as steric considerations and efficiency of binding.

A mouse PXR polypeptide is provided, and one or more mutations to be introduced into the mouse PXR amino acid sequence of the mouse PXR polypeptide are selected. The mutations are preferably selected so as to alter the mouse PXR polypeptide to be similar to a human PXR polypeptide. Although any mutations can be selected, preferred mutations are selected from the group consisting of an arginine to leucine substitution at residue 203 of the mouse PXR polypeptide, a protein to serine substitution at residue 205 of the mouse PXR polypeptide, a glutamine to histidine substitution at residue 404 of the mouse PXR polypeptide and a glutamine to arginine substitution at residue 407 of the mouse PXR polypeptide.

The selected one or more mutations is/are introduced into a mouse PXR polypeptide and a mutant mouse PXR polypeptide comprising the

selected one or more mutations selected is provided. The mutant mouse PXR polypeptide is expressed and purified using methods that are disclosed herein and will be apparent to those of skill in the art, upon consideration of the present disclosure.

The effect of the candidate therapeutic agent on the mutant mouse PXR polypeptide is determined. The effect can be determined by employing any appropriate reporter system. For example, the biological activity of a PXR polypeptide can be employed to determine an interaction.

The potential of the candidate therapeutic agent for use in humans is evaluated based on the effect of the candidate therapeutic agent on the mutant mouse PXR polypeptide. Thus, in one aspect of the present invention, the suitability of a candidate therapeutic for human disorders and conditions is conveniently assessed in a mouse model system. This system offers the ability to assess the biological response of a candidate therapeutic *in vivo*, which can add an additional degree of confidence to a conclusion made regarding the candidate therapeutic.

X. The Role of the Three-Dimensional Structure of the hPXR LDB in Solving Additional PXR Crystals

Because polypeptides can crystallize in more than one crystal form, the structural coordinates of a hPXR LBD, or portions thereof, as provided by the present invention, are particularly useful in solving the structure of other crystal forms of hPXR and the crystalline forms of other PXRs. The coordinates provided in the present invention can also be used to solve the structure of PXR or PXR LBD mutants (such as those described in Section VIII above), PXR LDB co-complexes, or of the crystalline form of any other protein with significant amino acid sequence homology to any functional domain of PXR.

IX.A. Determining the Three-Dimensional Structure of a Polypeptide Using the Three-Dimensional Structure of the hPXR LBD as a Template in Molecular Replacement

One method that can be employed for the purpose of solving additional PXR crystal structures is molecular replacement. See generally, Rossmann, ed, (1972) The Molecular Replacement Method, Gordon & Breach, New York. In the molecular replacement method, the unknown crystal structure, whether it is another crystal form of a PXR or a PXR LBD, (i.e. a PXR or a PXR LBD mutant), or a PXR or a PXR LBD polypeptide complexed with another compound (a "co-complex"), or the crystal of some other protein with significant amino acid sequence homology to any functional region of the a hPXR LBD, can be determined using the hPXR LBD structure coordinates provided in Table 4. This method provides an accurate structural form for the unknown crystal more quickly and efficiently than attempting to determine such information *ab initio*.

In addition, in accordance with this invention, PXR or PXR mutants (preferably hPXR or hPXR LBD mutants) can be crystallized in complex with known modulators. The crystal structures of a series of such complexes can then be solved by molecular replacement and compared with that of wild-type hPXR or the wild-type hPXR LBD. Potential sites for modification within the various binding sites of the enzyme can thus be identified. This information provides an additional tool for determining the most efficient binding interactions, for example, increased hydrophobic interactions, between the hPXR LBD and a chemical entity or compound.

All of the complexes referred to in the present disclosure can be studied using X-ray diffraction techniques (See, e.g., Blundell & Johnson (1985) *Method.Enzymol.*, 114A & 115B, (Wyckoff et al., eds.), Academic Press) and can be refined using computer software, such as the X-PLOR™ program (Brünger, (1992) *X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, Connecticut; X-PLOR is available from Molecular Simulations, Inc., San Diego, California). This information can thus be used to optimize known classes of PXR and PXR LBD modulators, and more importantly, to design and synthesize novel classes of PXR and PXR LBD modulators.

Laboratory Examples

The following Laboratory Examples have been included to illustrate preferred modes of the invention. Certain aspects of the following Laboratory Examples are described in terms of techniques and procedures found or contemplated by the present inventors to work well in the practice of the invention. These Laboratory Examples are exemplified through the use of standard laboratory practices of the inventors. In light of the present disclosure and the general level of skill in the art, those of skill will appreciate that the following Laboratory Examples are intended to be exemplary only and that numerous changes, modifications and alterations can be employed without departing from the spirit and scope of the invention.

Laboratory Example 1

PXR Expression and Purification

The PXR LBD expression construct was engineered as an N-terminal polyhistidine tagged fusion protein with residues 130-434 from the human PXR. The fusion insert (SEQ ID NO: 10) was subcloned into the pRSETA expression vector (Invitrogen, Carlsbad, California). DNA encoding residues 623-710 of the human SRC-1 gene (Onate et al., (1995) *Science* 270: 1354-57) (SEQ ID NO: 12) were subcloned into the bacterial vector pACYC184 (American Type Culture Collection #37033) along with a T7 promoter (Nolte et al., (1998) *Nature* 395: 137-43). The hPXRLBD/pRSETA and the SRC-1/pACYCl84 plasmids were co-transformed into the BL21(DE3) strain *of E. coli.* One-liter shake flask liquid cultures containing standard Luria-Bertani (LB) broth with 0.05 mg/ml ampicillin and 0.05 mg/ml chloramphenicol were inoculated and grown at 22°C for 20 hours. The cells were harvested by centrifugation (20 minutes, 3500g, 4°C) and the cell pellet was stored at -80°C. The cell pellet was resuspended in 250 ml of Buffer A (50 mM Tris-Cl pH 7.8, 250 mM NaCL, 50 mM imidazole pH 7.5, 5% glycerol). Cells were sonicated for 3-5 minutes on ice and the cell debris was removed by centrifugation (45 minutes, 20,000g, 4°C). The cleared supernatant was filtered through a 0.45 μM filter and loaded on to a 50 ml PROBOND™ nickel-

chelating resin (Invitrogen, Carlsbad, California). After washing to baseline with Buffer A, the column was washed with Buffer A containing 125 mM, imidazole pH 7.5. The PXR-LBD/SRC-1 complex was eluted from the column using Buffer A with 500 mM imidazole pH 7.5. Column fractions were pooled and concentrated using CENTRI-PREP™ 30K (Amicon/Millipore, Bedford, Massachusetts) units. The protein was subjected to size exclusion, using a column (26 mm X 90 cm) packed with SEPHAROSE™ S-75 resin (Amersham Pharmacia Biotech, Piscataway, New Jersey) pre-equilibrated with 20mM Tris-Cl pH 7.8, 250 mM NaCl, 5 mM DTT, 2.5 mM EDTA pH 8.0,5% glycerol. Column fractions containing the PXR/SRC-1 complex were pooled and diluted five fold with the dilution buffer (20 mM Tris-Cl pH 7.8, 5rnM DTT, 2.5mM EDTA pH 8.0, 5% glycerol) and loaded on to a 50 ml Q SEPHAROSE™ FAST FLOW™ column (Amersham Pharmacia Biotech, Piscataway, New Jersey). The column was washed to baseline with 20mM Tris-Cl pH 7.8,50 mM NaCL, 5 rnM DTT, 2.5 mM EDTA pH 8.0, 5% glycerol, and the protein complex was eluted from the column utilizing an increased salt gradient. The appropriate column fractions were pooled and stored on ice.

Laboratory Example 2

Crystallization of a PXR Ligand binding Domain

The human PXR ligand binding domain/SRC-1 complex (hPXR-LBD/SRC-1) was concentrated in the presence of 10-fold molar excesses of the SR12813 or I-SR12813 compounds to final concentrations of 4 and 5 mg/mL, respectively. The apo complex was concentrated to 5 mg/mL. Crystallization of hPXR-LBD was achieved by hanging-drop vapor diffusion against the following conditions at 22°C: SRI2813: 50 mM imidazole at pH 7.2,10% 2-propanol (v/v); I-SRI2813: 50 mM imidazole at pH 7.2, 11% 2-propanol (v/v); apo: 50 mM imidazole at pH 7.4,10% 2-propanol (v/v).

Laboratory Example 3

Structure Determination

The structures of an apo (unliganded) and SRI28I3-bound form of the ligand binding domain of human PXR were determined by molecular replacement using the crystal structure of the vitamin D receptor (VDR) as search models (SEQ ID NO: 7, structural coordinates presented in Table 5). Non-identical side chains were trimmed prior to rotation and translation function searches in AmoRe (Navaza & Saludjian, (1997) *Method Enzymol.* 276A: 581-94); clear solutions were obtained for each in the proper enantiomorphic space group, $P4_32_12$. The structures were refined using the torsion angle protocol in CNS with the maximum likelihood function as a target, and included an overall anisotropic B-factor and a bulk solvent correction (Brünger et al., (1998) *Acta Crystallog. D* 54: 905-21).

For both structures, 10% of the observed data were set aside for cross-validation using the free-R statistic prior to any structural refinement (Brünger, (1993) *Acta Crystallog. D* 49: 24-36). Manual adjustments and rebuilding of the model were performed using O (Jones et al., (1991) *Acta Crystallog. A* 47:110-19) and σA-weighted electron density maps (Read, (1986) *Acta Crystallog. A* 42: 140-49). At the later stages of refinement 281 and 166 period solvent sites were added to the apo and SR12813 complexes, respectively. Structures exhibit good geometry with no Ramachandran outliers, as shown in Table 1. Molecular graphics figures were created with MOLSCRIPT (Kraulis, (1991) *J. Applied Crystallogr.* 24: 946-50), RASTER-3D (Merritt & Bacon, (1997) *Method Enzymol.* 277: 505-24), GRASP (Nicholls et al., (1991) *Proteins* 11: 281-96) and CS CHEMDRAW PRO™, available from CambridgeSoft of Cambridge, Massachusetts.

Laboratory Example 4

Positioning of SR12813 Ligands

Clear electron density was evident in the ligand binding cavity of PXR at the earliest stages of refinement. The positioning of single SR12813 molecules into this electron density, however, resulted in significant positive

electron density peaks elsewhere in the cavity, suggesting that more than one conformation of the ligand was present. This was confirmed using difference density generated from crystals of hPXR-LBD containing an iodinated form of SR12813 (I-SR12813). The two t-butyl groups of SR12813 were replaced with iodine in I-SR12813; this ligand was found to activate PXR as well as the parent SR12813 compound. Two difference density maps were used to guide the positioning of SR12813 ligands: [(| $F_{obs}^{ISR12813}$ | - | $F_{obs}^{SR12813}$ |), $\phi_{calc}^{PROTEIN}$] and [(| $F_{obs}^{ISR12813}$ | - | $F_{obs}^{PROTEIN}$ |), $\phi_{calc}^{PROTEIN}$].

Each map was calculated to 3.0 Å resolution and both showed a consistent set of significant positive difference density peaks in the ligand binding cavity, which were interpreted to be the positions of the iodine atoms on ISR12813. These peaks were the highest difference density peaks in both maps. It was found that three SR12813 ligands optimally satisfied the electron density in the ligand binding cavity. The first SR12813 (position 1) was placed in part based on two 7σ difference density peaks at appropriate positions in both maps. The second SR12813 (position 3) was positioned using two 5σ difference density peaks at appropriate positions. The third SR12813 (position 2) was placed based on 4σ iodine difference density peaks and residual difference density in the maps after the refinement of positions 1 and 3 together.

Standard simulated density maps, [(| $F_{obs}^{PROTEIN}$ | - | $F_{calc}^{PROTEIN}$ |), $\phi_{calc}^{PROTEIN}$], at 2.75 Å resolution were used to guide the positioning of the remainder of each SR12813 ligand and to assess the quality of the fits to the density (Brünger et al., (1990) *Acta Crystallog. A* 46: 585-93). No positive or negative difference density peaks above 30' were present in the binding cavity after the refinement of the three ligand positions together. Occupancies of each ligand position were estimated based on iodine difference density and standard electron density peaks, and were fixed at 0.5, 0.2 and 0.3 for positions 1, 2 and 3, respectively. In addition, the R and R-free after the positioning of the three ligand positions were improved over each conformation alone, or over pairs of each conformation.

Laboratory Example 5

Mutagenesis

The human and mouse PXR mutants were generated using the QUIKCHANGE™ mutagenesis kit (Stratagene, La Jolla, California) according to the manufacturer's instructions. All mutants were confirmed by sequence analysis.

Laboratory Example 6

Cotransfection Assays

Transient cotransfection experiments were performed using CV -1 cells and the XREM-CYP3A4-LUC reporter, containing the enhancer and promoter of CYP3A4 driving luciferase expression, as described in Moore et al., (2000) *J. Biol. Chem.* 275: 15122-27.

References

The references listed below as well as all references cited in the specification are incorporated herein by reference to the extent that they supplement, explain, provide a background for or teach methodology, techniques and/or compositions employed herein.

Altschul et al., (1990) *J. Mol. Biol.* 215: 403-10

Ausubel et al., (1989) Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, New York

Bartlett et al., (1989) *Special Pub., Royal Chem. Soc.* 78: 182-96

Berkhout et al., (1996) *J. Biol. Chem.* 271: 14376-82

Berkhout et al., (1997) *Atherosclerosis* 133: 203-21

Bertilsson et al., (1998) *Proc. Nat. Acad. Sci. U.S.A.* 95: 12208-13

Blumberg et al., (1998) *Gene Dev.* 12: 3195-205

Blundell et al., (1976) Protein Crystallography, Academic Press

Blundell & Johnson, (1985) *Method.Enzymol.*, 114A & 115B, (Wyckoff et al., eds.), Academic Press

Bohm, (1992) *J. Comput. Aid. Mol. Des.*, 6: 61-78

Bondi, (1964) *J. Phys. Chem.* 68: 441-51

Bourget et al., (1995) *Nature* 375: 377-82

Brooks et al., (1983) *J. Comp. Chem.*, 8: 132

Brünger et al., (1990) *Acta Crystallog. A* 46: 585-93

Brünger, (1992) *X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, Connecticut

Brünger, (1993) *Acta Crystallog. D* 49: 24-36

Brünger et al., (1998) *Acta Crystallog. D* 54: 905-21

Brzozowski et al., (1997) *Nature (London)* 389: 753-58

Case et al., (1997), AMBER 5, University of California, San Francisco

Cohen et al., (1990) *J. Med. Chem.* 33: 883-94

Creighton, (1983) Proteins: Structures and Molecular Principles, W.H. Freeman & Co., New York

Drewes et al., (1996) *Mol. Cell. Biol.* 16:925-31

Ducruix & Geige, (1992) Crystallization of Nucleic Acids and Proteins: A Practical Approach, IRL Press, Oxford, England

Eisen et al., (1994) *Proteins* 19: 199-221

Gampe, Jr. et al., (2000) *Mol. Cell* 5: 545-55

Goodford, (1985) *J. Med. Chem.* 28: 849-57

Goodsell & Olsen, (1990) *Proteins* 8: 195-202

Goodwin et al., (1999) *Mol. Pharmacol.* 56: 1329-39

Gribskov et al., (1986) *Nucl. Acids. Res.* 14: 6745

Hauptman, (1997) *Curr. Opin. Struct. Biol.* 7: 672-80

Jones et al., (1991) *Acta Crystallog. A* 47:110-19

Jones et al., (2000) *Mol. Endocrinol.* 14: 27-39

Kliewer et al., (1998) *Cell* 92: 730-82

Kraulis, (1991) *J. Applied Crystallogr.* 24: 946-50

Kuntz et al., (1992) *J. Mol. Biol.* 161: 269-88

Kyte & Doolittle, (1982), *J. Mol. Biol.* 157: 105-132

Lambert, (1997) in Practical Application of Computer-Aided Drug Design, (Charifson, ed.) Marcel-Dekker, New York, pp. 243-303

Lattman, (1985) *Method Enzymol.*, 115: 55-77

Lehmann et al., (1999) *J. Clin. Invest.* 102: 1016-23

Martin, (1992) *J. Med. Chem.* 35: 2145-54

Maurel, (1996) in Cytochrome P450: Metabolic and Toxicological Aspects (Ionnides, ed.) CRC Press, Inc., Boca Raton, Florida, 241-70

McPherson, (1982) Preparation and Analysis of Protein Crystals, John Wiley, New York

McPherson, (1990) *Eur. J. Biochem.* 189:1-23

McPherson et al., (1989) Preparation and Analysis of Protein Crystals, Robert E. Krieger Publishing Company, Malabar, Florida

McRee, (1992) *J. Mol. Graphics* 10: 44-47

Merritt & Bacon, (1997) *Method Enzymol.* 277: 505-24

Miranker & Karplus, (1991) *Proteins* 11: 29-34

Moore et al., (2000) *J. Biol. Chem.* 275: 15122-27

Moras & Gronenmeyer, (1998) *Curr. Opin. Cell Biol.* 10: 384-91

Navaza & Saludjian, (1997) *Method Enzymol.* 276A: 581-94

Navia & Murcko, (1992) *Curr. Opin. Struc. Biol.* 2: 202-10

Needleman et al., (1970) *J. Mol. Biol.* 48: 443

Nicholls et al., (1991) *Proteins* 11: 281-96

Nishibata & Itai, (1991) *Tetrahedron* 47: 8985

Nolte et al., (1998) *Nature (London)* 395: 137-43

Onate et al., (1995) *Science* 270: 1354-57

Otwinowski, (1991), in Isomorphous Replacement and Anomalous Scattering, (Evans & Leslie, eds.), pp. 80-86, Daresbury Laboratory, Daresbury, United Kingdom

Pearlman et al., (1995) *Comput. Phys. Commun.* 91: 1-41

Piscitelli et al., (2000) *Lancet* 355: 547-48

Rarey et al., (1996) *J. Comput. Aid. Mol. Des.* 10:41-54

Read, (1986) *Acta Crystallog. A* 42: 140-49

Renaud et al., (1995) *Nature* 378: 681-89

Rochel et al., *Eur. J. Biochem.* 268: 971-79

Rochel et al., (2000) *Mol. Cell* 5: 173-79

Rossmann, ed, (1972) The Molecular Replacement Method, Gordon & Breach, New York

Ruschitzka et al., (2000) *Lancet* 355: 548-49

Samanta et al., (1999) *Acta Crystalog. D* 55: 1421-27

Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York

Savas et al., (1999) *Mol. Pharmacol.* 56: 851-57

Savas et al., (2000) *Drug Metab Dispos.* 28: 529-37

Schwartz et al., eds., (1979), Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, pp. 357-358

Sheldrick, (1990) *Acta Cryst.* A46: 467

Shiau et al., (1998) *Cell* 95: 927-37

Sladek et al., *Genes Dev.* 4: 2353-65

Smith et al., (1981) *Adv. Appl. Math.* 2: 482

Van Holde, (1971) Physical Biochemistry, Prentice-Hall, N. J., 221-39

Waxman, (1999) *Arch. Biochem Biophys.* 369: 11-23

Weatherman et al., (1999) *Annu. Rev. Biochem.* 68: 559-81

Weber, (1991) *Adv. Protein Chem.* 41:1-36

Weeks et al., (1993) *Acta Cryst.* D49: 179

Wellner, (1971) *Anal. Chem.* 43: 597

Wentworth et al., (2000) *J. Endocrinol.* 166: R11-R16

Wetmur & Davidson, (1968) *J. Mol. Biol.* 31: 349-70

Williams & Sigler, (1998) *Nature* 393: 392-96

Xie et al., (2000) *Nature* 406: 435-39

Xu et al., (1999) *Mol. Cell* 3: 397-403

U.S. Patent No. 4,554,101

U.S. Patent No. 4,672,108

U.S. Patent No. 5,463,564

U.S. Patent No. 5,834,228

U.S. Patent No. 5,872,011

U.S. Patent No. 6,008,033

WO 84/03564

WO 2000/037077

WO 2000/025134

TABLE 1

CRYSTALLOGRAPHIC DATA AND REFINEMENT

	APO	SR12813	I-SR12813[∞]
Resolution (Å; highest shell)	20-2.5 (2.54-2.5)	20-2.75 (2.8-2.75)	20-3.0 (3.05-3.0)
Space Group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell Constants (Å)	A=91.6; c=85.0	a=91.3; c=85.3	a=91.1; c=84.5
Total Reflections	179,671	133,315	91,752
Unique Reflections	12,611	9,512	7,269
Mean Redundancy	14.2	14	12.6
R_{sym}* (%; highest shell)	7.3 (39.0)	9.8 (40.0)	9.7 (51.0)
Completeness (%; highest shell)	99.9 (100)	99.6 (100)	99.9 (100)
Mean I/σ (highest shell)	40.8 (7.9)	25.9 (3.7)	31.4 (5.4)
R_{cryst}† (%; highest shell)	20.9 (25.9)	21.3 (26.1)	--
R_{free}‡ (%; highest shell)	28.4 (32.1)	27.4 (40.5)	--
RMSD§ Bond Lengths (Å)	0.0059	0.0093	--
RMSD§ Bond Angles (°)	1.11	1.29	--
RMSD§ Dihedrals (°)	20.54	21.33	--
RMSD§ Impropers (°)	0.71	0.91	--
Number Protein Atoms	2141	2141	--
Number Solvent Sites	281	166	--
Number Ligand Atoms	--	99	--

[∞] Data from crystals containing I-SR12813 were used only to guide and confirm the positioning of SR12813 ligands; thus, a refined structure was not generated.

* $R_{sym} = \Sigma|I-\langle I\rangle| / \Sigma I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity of multiple symmetry-related observations of that reflection.

† $R_{cryst} = \Sigma||F_{obs}|-|F_{calc}|| / \Sigma|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.

‡ $R_{free} = \Sigma||F_{obs}|-|F_{calc}|| / \Sigma|F_{obs}|$ for 10% of the data not used at any stage of structural refinement.

§ RMSD, root mean square deviation.

TABLE 2

COMPARISON OF HUMAN PXR-LBD WITH NUCLEAR RECEPTOR LBDs OF KNOWN STRUCTURE

	Number Residues Superimposed	Sequence Identity (%)	RMSD (Å)	Ligand binding Cavity Volume* ($Å^3$)
PXR	--	--	--	1,150
Vitamin D Receptor	225	45	1.8	871
RXRα	178	22	2.8	687
PPARγ	203	25	2.6	1,619
Progesterone Receptor	201	19	2.9	557
Estrogen Receptor	180	18	2.8	476

* The cavity volumes were calculated with GRASP (Nicholls et al., (1991) *Proteins* 11: 281-96) using a grid spacing of 0.2 Å, probe radius of 1.4 Å and the atomic radii of Bondi (H=1.2 Å, C=1.7 Å, N=1.55 Å, O=1.52 Å, S=1.8 Å) (Bondi, (1964) *J. Phys. Chem.* 68: 441-51), where hydrogens were treated explicitly. Openings in the PXR, PPARγ and RXRα pockets were closed by covering these proteins with an external layer of water molecules prior to calculation of the molecular surface and cavities.

TABLE 3

CALCULATED BURIED SOLVENT-ACCESSIBLE SURFACE AREA OF AMINO ACID RESIDUES THAT LINE THE LIGAND BINDING CAVITY OF HPXR

Residue	Buried Surface ($Å^2$)	Contacts SR12813 Ligand?	Human	Rabbit	Rat	Mouse
206	9.8	*Yes*	L	L	R	S
208	15.8	*Yes*	S	T	P	P
209	30.5	*Yes*	*L*	*M*	*M*	*M*
211	16.3	*Yes*	*V*	*L*	*I*	*I*
240	14.8	*Yes*	**L**	**L**	**L**	**L**
243	43.6	*Yes*	*M*	*L*	*L*	*L*
244	1.9	*No*	**A**	**A**	**A**	**A**
246	6.1	*Yes*	*M*	*M*	*V*	*V*
247	30.4	*Yes*	**S**	**S**	**S**	**S**
251	11.2	*Yes*	**F**	**F**	**F**	**F**
281	29.3	*Yes*	*F*	*L*	*F*	*F*
284	9.6	*Yes*	**C**	**C**	**C**	**C**
285	23.0	*Yes*	Q	L	I	I
288	29.6	*Yes*	**F**	**F**	**F**	**F**
299	**28.6**	*Yes*	**W**	**W**	**W**	**W**
306	12.9	*No*	**Y**	**Y**	**Y**	**Y**
308	3.7	*No*	*L*	*V*	*F*	*F*
321	4.2	*No*	*E*	*D*	*D*	*D*
323	20.3	*Yes*	*M*	*L*	*L*	*L*
324	12.0	*Yes*	*L*	*L*	*M*	*M*
327	2.5	*No*	**H**	**H**	**H**	**H**
407	41.6	*Yes*	H	H	Q	Q
410	14.7	*Yes*	R	R	Q	Q
411	33.1	*No*	**L**	**L**	**L**	**L**
414	11.5	*No*	**I**	**I**	**I**	**I**
420	9.3	*Yes*	**F**	**F**	**F**	**F**
425	6.2	*No*	**M**	**M**	**M**	**M**
429	3.6	*No*	**F**	**F**	**F**	**F**

Boldface indicates identity; italics indicate similarity. The buried surface area upon ligand binding was calculated using CNS (Brünger et al., (1998) *Acta Crystallogr. D* 54: 905-21). The four residues underlined were mutated to examine the ligand specificity of mouse and human PXR.

TABLE 4

ATOMIC COORDINATE DATA FOR CRYSTALLIZED HUMAN PXR

ATOM	ATOM TYPE	RESIDUE	#	X	Y	Z	OCC	B
1	C	GLY	142	-5.808	44.753	13.561	1.00	58.97
2	O	GLY	142	-5.723	45.523	14.515	1.00	59.54
3	N	GLY	142	-4.377	43.177	14.842	1.00	59.37
4	CA	GLY	142	-5.307	43.330	13.685	1.00	59.68
5	N	LEU	143	-6.324	45.108	12.387	1.00	58.87
6	CA	LEU	143	-6.839	46.455	12.152	1.00	58.50
7	CB	LEU	143	-6.483	46.907	10.736	1.00	57.90
8	CG	LEU	143	-5.849	48.290	10.555	1.00	57.77
9	CD1	LEU	143	-4.599	48.411	11.407	1.00	56.51
10	CD2	LEU	143	-5.505	48.492	9.090	1.00	56.92
11	C	LEU	143	-8.352	46.446	12.333	1.00	58.92
12	O	LEU	143	-9.046	45.640	11.714	1.00	59.85
13	N	THR	144	-8.862	47.341	13.174	1.00	58.88
14	CA	THR	144	-10.299	47.407	13.444	1.00	59.76
15	CB	THR	144	-10.594	48.134	14.770	1.00	58.88
16	OG1	THR	144	-10.447	49.549	14.585	1.00	56.11
17	CG2	THR	144	-9.644	47.660	15.857	1.00	58.12
18	C	THR	144	-11.127	48.103	12.366	1.00	61.28
19	O	THR	144	-10.604	48.831	11.518	1.00	62.11
20	N	GLU	145	-12.436	47.875	12.419	1.00	62.26
21	CA	GLU	145	-13.360	48.482	11.477	1.00	62.32
22	CB	GLU	145	-14.799	48.130	11.870	1.00	63.47
23	CG	GLU	145	-15.885	48.842	11.066	1.00	67.03
24	CD	GLU	145	-15.765	48.632	9.560	1.00	68.86
25	OE1	GLU	145	-14.952	47.780	9.129	1.00	70.24
26	OE2	GLU	145	-16.492	49.322	8.809	1.00	68.98
27	C	GLU	145	-13.163	49.994	11.476	1.00	61.32
28	O	GLU	145	-12.997	50.600	10.424	1.00	60.90
29	N	GLU	146	-13.169	50.594	12.662	1.00	60.97
30	CA	GLU	146	-12.988	52.037	12.793	1.00	61.06
31	CB	GLU	146	-13.026	52.452	14.269	1.00	64.21
32	CG	GLU	146	-12.567	51.378	15.254	1.00	68.91
33	CD	GLU	146	-13.638	50.326	15.526	1.00	71.06
34	OE1	GLU	146	-14.703	50.686	16.079	1.00	71.24
35	OE2	GLU	146	-13.415	49.142	15.184	1.00	72.11
36	C	GLU	146	-11.688	52.514	12.161	1.00	59.14
37	O	GLU	146	-11.652	53.556	11.504	1.00	59.31
38	N	GLN	147	-10.620	51.754	12.371	1.00	57.26
39	CA	GLN	147	-9.325	52.095	11.802	1.00	54.84
40	CB	GLN	147	-8.248	51.126	12.309	1.00	54.51
41	CG	GLN	147	-7.812	51.390	13.755	1.00	53.99
42	CD	GLN	147	-6.957	50.275	14.353	1.00	54.37
43	OE1	GLN	147	-6.291	50.474	15.368	1.00	55.65
44	NE2	GLN	147	-6.984	49.096	13.737	1.00	54.19
45	C	GLN	147	-9.430	52.034	10.283	1.00	53.65
46	O	GLN	147	-8.948	52.921	9.581	1.00	52.79
47	N	ARG	148	-10.083	50.991	9.783	1.00	52.52
48	CA	ARG	148	-10.259	50.819	8.349	1.00	51.82

49	CB	ARG	148	-11.020	49.528	8.061	1.00	52.78
50	CG	ARG	148	-10.368	48.278	8.626	1.00	55.07
51	CD	ARG	148	-11.157	47.046	8.233	1.00	56.61
52	NE	ARG	148	-11.103	46.818	6.792	1.00	59.42
53	CZ	ARG	148	-11.976	46.074	6.116	1.00	61.30
54	NH1	ARG	148	-12.982	45.485	6.752	1.00	62.10
55	NH2	ARG	148	-11.839	45.912	4.805	1.00	62.11
56	C	ARG	148	-11.010	51.991	7.732	1.00	50.99
57	O	ARG	148	-10.585	52.545	6.720	1.00	52.39
58	N	MET	149	-12.129	52.369	8.336	1.00	50.18
59	CA	MET	149	-12.920	53.472	7.813	1.00	50.13
60	CB	MET	149	-14.194	53.665	8.643	1.00	53.54
61	CG	MET	149	-15.135	52.461	8.644	1.00	57.84
62	SD	MET	149	-16.690	52.746	9.560	1.00	64.69
63	CE	MET	149	-16.093	52.876	11.284	1.00	59.97
64	C	MET	149	-12.105	54.755	7.821	1.00	48.09
65	O	MET	149	-12.210	55.576	6.912	1.00	48.07
66	N	MET	150	-11.285	54.918	8.850	1.00	45.74
67	CA	MET	150	-10.458	56.107	8.990	1.00	43.72
68	CB	MET	150	-9.721	56.049	10.318	1.00	43.70
69	CG	MET	150	-9.356	57.394	10.877	1.00	45.87
70	SD	MET	150	-8.499	57.190	12.442	1.00	49.99
71	CE	MET	150	-9.883	56.803	13.553	1.00	48.07
72	C	MET	150	-9.453	56.229	7.845	1.00	42.57
73	O	MET	150	-9.258	57.307	7.275	1.00	41.08
74	N	ILE	151	-8.822	55.111	7.507	1.00	41.43
75	CA	ILE	151	-7.837	55.087	6.439	1.00	40.06
76	CB	ILE	151	-7.118	53.726	6.387	1.00	39.91
77	CG2	ILE	151	-6.064	53.727	5.273	1.00	39.50
78	CG1	ILE	151	-6.449	53.459	7.742	1.00	38.84
79	CD1	ILE	151	-5.681	52.170	7.818	1.00	37.16
80	C	ILE	151	-8.503	55.377	5.105	1.00	39.71
81	O	ILE	151	-8.014	56.190	4.311	1.00	39.37
82	N	ARG	152	-9.627	54.717	4.864	1.00	39.09
83	CA	ARG	152	-10.363	54.920	3.627	1.00	37.68
84	CB	ARG	152	-11.571	53.980	3.584	1.00	40.07
85	CG	ARG	152	-12.593	54.281	2.493	1.00	45.57
86	CD	ARG	152	-11.962	54.439	1.111	1.00	48.14
87	NE	ARG	152	-12.786	53.812	0.080	1.00	50.92
88	CZ	ARG	152	-12.837	52.500	-0.134	1.00	50.95
89	NH1	ARG	152	-12.106	51.678	0.608	1.00	50.25
90	NH2	ARG	152	-13.623	52.006	-1.082	1.00	50.40
91	C	ARG	152	-10.807	56.373	3.491	1.00	35.71
92	O	ARG	152	-10.785	56.926	2.399	1.00	34.62
93	N	GLU	153	-11.203	56.996	4.595	1.00	35.31
94	CA	GLU	153	-11.641	58.384	4.537	1.00	36.30
95	CB	GLU	153	-12.264	58.822	5.864	1.00	39.33
96	CG	GLU	153	-12.848	60.228	5.808	1.00	45.70
97	CD	GLU	153	-13.770	60.561	6.982	1.00	52.48
98	OE1	GLU	153	-14.304	61.695	6.992	1.00	54.21
99	OE2	GLU	153	-13.971	59.708	7.888	1.00	53.74
100	C	GLU	153	-10.459	59.278	4.202	1.00	35.22
101	O	GLU	153	-10.594	60.257	3.460	1.00	33.48
102	N	LEU	154	-9.297	58.938	4.753	1.00	34.10
103	CA	LEU	154	-8.089	59.706	4.486	1.00	32.93

104	CB	LEU	154	-6.973	59.284	5.436	1.00	32.59
105	CG	LEU	154	-7.114	59.676	6.903	1.00	32.45
106	CD1	LEU	154	-5.829	59.293	7.648	1.00	28.94
107	CD2	LEU	154	-7.391	61.180	7.007	1.00	30.74
108	C	LEU	154	-7.648	59.486	3.040	1.00	32.49
109	O	LEU	154	-7.273	60.438	2.342	1.00	31.27
110	N	MET	155	-7.689	58.227	2.596	1.00	31.58
111	CA	MET	155	-7.307	57.877	1.227	1.00	31.69
112	CB	MET	155	-7.271	56.361	1.046	1.00	28.61
113	CG	MET	155	-6.056	55.708	1.673	1.00	29.23
114	SD	MET	155	-4.498	56.482	1.116	1.00	28.89
115	CE	MET	155	-4.096	55.468	-0.360	1.00	30.07
116	C	MET	155	-8.256	58.485	0.200	1.00	32.29
117	O	MET	155	-7.836	58.922	-0.871	1.00	32.41
118	N	ASP	156	-9.539	58.518	0.535	1.00	32.13
119	CA	ASP	156	-10.530	59.078	-0.367	1.00	32.52
120	CB	ASP	156	-11.938	58.791	0.172	1.00	35.14
121	CG	ASP	156	-13.033	59.420	-0.671	1.00	36.51
122	OD1	ASP	156	-13.515	60.517	-0.303	1.00	35.94
123	OD2	ASP	156	-13.404	58.814	-1.702	1.00	39.04
124	C	ASP	156	-10.287	60.575	-0.496	1.00	31.62
125	O	ASP	156	-10.373	61.140	-1.586	1.00	30.81
126	N	ALA	157	-9.960	61.213	0.621	1.00	31.28
127	CA	ALA	157	-9.700	62.646	0.628	1.00	30.89
128	CB	ALA	157	-9.547	63.128	2.050	1.00	29.70
129	C	ALA	157	-8.448	62.977	-0.182	1.00	32.21
130	O	ALA	157	-8.377	64.029	-0.815	1.00	31.78
131	N	GLN	158	-7.469	62.070	-0.155	1.00	32.47
132	CA	GLN	158	-6.216	62.245	-0.880	1.00	33.24
133	CB	GLN	158	-5.214	61.153	-0.497	1.00	35.65
134	CG	GLN	158	-4.160	61.542	0.535	1.00	37.85
135	CD	GLN	158	-3.119	62.520	0.020	1.00	34.82
136	OE1	GLN	158	-3.399	63.703	-0.168	1.00	35.15
137	NE2	GLN	158	-1.907	62.028	-0.203	1.00	34.05
138	C	GLN	158	-6.380	62.215	-2.392	1.00	34.09
139	O	GLN	158	-5.916	63.114	-3.085	1.00	34.76
140	N	MET	159	-7.019	61.173	-2.913	1.00	34.17
141	CA	MET	159	-7.176	61.089	-4.352	1.00	34.32
142	CB	MET	159	-7.639	59.686	-4.764	1.00	35.03
143	CG	MET	159	-9.088	59.365	-4.524	1.00	37.89
144	SD	MET	159	-10.089	59.609	-6.008	1.00	40.60
145	CE	MET	159	-11.625	60.101	-5.262	1.00	36.87
146	C	MET	159	-8.114	62.167	-4.885	1.00	33.55
147	O	MET	159	-8.042	62.537	-6.049	1.00	35.81
148	N	LYS	160	-8.979	62.697	-4.038	1.00	31.63
149	CA	LYS	160	-9.872	63.743	-4.487	1.00	31.37
150	CB	LYS	160	-11.075	63.874	-3.547	1.00	32.46
151	CG	LYS	160	-12.260	62.968	-3.879	1.00	34.88
152	CD	LYS	160	-13.332	63.011	-2.786	1.00	35.55
153	CE	LYS	160	-14.502	62.082	-3.107	1.00	37.74
154	NZ	LYS	160	-15.517	62.040	-2.004	1.00	39.58
155	C	LYS	160	-9.159	65.088	-4.567	1.00	31.90
156	O	LYS	160	-9.495	65.912	-5.420	1.00	33.35
157	N	THR	161	-8.166	65.310	-3.703	1.00	31.15
158	CA	THR	161	-7.460	66.596	-3.657	1.00	29.94

159	CB	THR	161	-7.559	67.211	-2.244	1.00	29.42
160	OG1	THR	161	-6.842	66.388	-1.317	1.00	29.69
161	CG2	THR	161	-9.009	67.303	-1.801	1.00	26.99
162	C	THR	161	-5.983	66.645	-4.071	1.00	30.48
163	O	THR	161	-5.371	67.716	-4.086	1.00	31.55
164	N	PHE	162	-5.402	65.505	-4.413	1.00	30.31
165	CA	PHE	162	-4.004	65.484	-4.809	1.00	29.27
166	CB	PHE	162	-3.259	64.420	-3.983	1.00	26.54
167	CG	PHE	162	-1.751	64.499	-4.072	1.00	24.19
168	CD1	PHE	162	-0.960	63.605	-3.347	1.00	21.93
169	CD2	PHE	162	-1.119	65.431	-4.899	1.00	21.10
170	CE1	PHE	162	0.444	63.634	-3.449	1.00	22.80
171	CE2	PHE	162	0.284	65.467	-5.009	1.00	22.24
172	CZ	PHE	162	1.067	64.567	-4.285	1.00	21.06
173	C	PHE	162	-3.886	65.205	-6.310	1.00	30.34
174	O	PHE	162	-4.004	64.066	-6.756	1.00	29.81
175	N	ASP	163	-3.660	66.268	-7.078	1.00	32.71
176	CA	ASP	163	-3.498	66.189	-8.535	1.00	34.12
177	CB	ASP	163	-3.788	67.575	-9.147	1.00	36.50
178	CG	ASP	163	-3.752	67.582	-10.679	1.00	39.50
179	OD1	ASP	163	-4.204	66.595	-11.304	1.00	39.89
180	OD2	ASP	163	-3.293	68.598	-11.257	1.00	39.21
181	C	ASP	163	-2.054	65.754	-8.820	1.00	33.62
182	O	ASP	163	-1.206	66.568	-9.184	1.00	33.72
183	N	THR	164	-1.785	64.464	-8.651	1.00	33.48
184	CA	THR	164	-0.445	63.923	-8.846	1.00	35.68
185	CB	THR	164	-0.414	62.407	-8.555	1.00	37.91
186	OG1	THR	164	-1.373	61.736	-9.384	1.00	40.43
187	CG2	THR	164	-0.737	62.148	-7.082	1.00	39.32
188	C	THR	164	0.208	64.176	-10.202	1.00	35.58
189	O	THR	164	1.406	63.957	-10.360	1.00	36.84
190	N	THR	165	-0.562	64.629	-11.181	1.00	35.30
191	CA	THR	165	0.003	64.915	-12.500	1.00	35.07
192	CB	THR	165	-0.896	64.358	-13.655	1.00	37.29
193	OG1	THR	165	-2.166	65.028	-13.655	1.00	37.40
194	CG2	THR	165	-1.122	62.851	-13.488	1.00	36.09
195	C	THR	165	0.149	66.433	-12.658	1.00	34.40
196	O	THR	165	0.716	66.923	-13.635	1.00	32.66
197	N	PHE	166	-0.377	67.175	-11.688	1.00	33.66
198	CA	PHE	166	-0.288	68.621	-11.725	1.00	34.14
199	CB	PHE	166	1.189	69.034	-11.680	1.00	33.49
200	CG	PHE	166	1.836	68.882	-10.321	1.00	29.99
201	CD1	PHE	166	2.016	69.990	-9.493	1.00	29.19
202	CD2	PHE	166	2.282	67.642	-9.877	1.00	29.01
203	CE1	PHE	166	2.637	69.861	-8.240	1.00	27.97
204	CE2	PHE	166	2.904	67.506	-8.626	1.00	27.66
205	CZ	PHE	166	3.081	68.618	-7.809	1.00	24.70
206	C	PHE	166	-0.954	69.169	-12.997	1.00	35.39
207	O	PHE	166	-0.582	70.231	-13.496	1.00	35.45
208	N	SER	167	-1.947	68.445	-13.506	1.00	36.79
209	CA	SER	167	-2.657	68.844	-14.718	1.00	38.60
210	CB	SER	167	-3.743	67.827	-15.064	1.00	39.08
211	OG	SER	167	-3.175	66.652	-15.611	1.00	43.62
212	C	SER	167	-3.292	70.219	-14.683	1.00	40.09
213	O	SER	167	-3.446	70.860	-15.729	1.00	40.68

214	N	HIS	168	-3.674	70.678	-13.497	1.00	41.18
215	CA	HIS	168	-4.308	71.986	-13.403	1.00	42.15
216	CB	HIS	168	-5.549	71.903	-12.528	1.00	44.04
217	CG	HIS	168	-6.527	70.882	-13.006	1.00	47.91
218	CD2	HIS	168	-7.674	71.008	-13.713	1.00	49.02
219	ND1	HIS	168	-6.324	69.530	-12.838	1.00	49.74
220	CE1	HIS	168	-7.304	68.866	-13.421	1.00	51.16
221	NE2	HIS	168	-8.136	69.739	-13.960	1.00	51.57
222	C	HIS	168	-3.395	73.081	-12.911	1.00	41.34
223	O	HIS	168	-3.856	74.121	-12.454	1.00	42.51
224	N	PHE	169	-2.096	72.843	-13.004	1.00	39.52
225	CA	PHE	169	-1.134	73.839	-12.593	1.00	38.30
226	CB	PHE	169	0.050	73.195	-11.868	1.00	34.93
227	CG	PHE	169	1.099	74.182	-11.439	1.00	30.28
228	CD1	PHE	169	0.747	75.316	-10.711	1.00	27.54
229	CD2	PHE	169	2.443	73.978	-11.761	1.00	28.72
230	CE1	PHE	169	1.723	76.237	-10.307	1.00	28.10
231	CE2	PHE	169	3.427	74.885	-11.366	1.00	25.29
232	CZ	PHE	169	3.066	76.020	-10.636	1.00	26.54
233	C	PHE	169	-0.657	74.506	-13.858	1.00	39.42
234	O	PHE	169	0.202	73.973	-14.554	1.00	41.07
235	N	LYS	170	-1.235	75.659	-14.173	1.00	41.29
236	CA	LYS	170	-0.852	76.403	-15.368	1.00	42.81
237	CB	LYS	170	-1.661	75.921	-16.574	1.00	42.89
238	CG	LYS	170	-3.155	75.723	-16.326	1.00	45.42
239	CD	LYS	170	-3.739	74.778	-17.387	1.00	47.25
240	CE	LYS	170	-5.236	74.530	-17.201	1.00	49.85
241	NZ	LYS	170	-6.090	75.701	-17.565	1.00	49.98
242	C	LYS	170	-1.015	77.906	-15.172	1.00	43.97
243	O	LYS	170	-1.596	78.354	-14.177	1.00	43.83
244	N	ASN	171	-0.483	78.685	-16.111	1.00	44.71
245	CA	ASN	171	-0.559	80.141	-16.020	1.00	44.49
246	CB	ASN	171	-2.014	80.587	-15.835	1.00	46.65
247	CG	ASN	171	-2.939	79.996	-16.878	1.00	48.97
248	OD1	ASN	171	-4.089	79.680	-16.584	1.00	50.99
249	ND2	ASN	171	-2.445	79.851	-18.107	1.00	48.49
250	C	ASN	171	0.262	80.588	-14.812	1.00	42.76
251	O	ASN	171	-0.141	81.490	-14.071	1.00	44.46
252	N	PHE	172	1.403	79.941	-14.609	1.00	38.82
253	CA	PHE	172	2.273	80.269	-13.491	1.00	34.81
254	CB	PHE	172	2.727	78.986	-12.773	1.00	33.64
255	CG	PHE	172	3.264	77.923	-13.688	1.00	29.85
256	CD1	PHE	172	4.535	78.031	-14.237	1.00	29.38
257	CD2	PHE	172	2.490	76.815	-14.011	1.00	29.76
258	CE1	PHE	172	5.033	77.048	-15.100	1.00	30.46
259	CE2	PHE	172	2.977	75.821	-14.873	1.00	31.08
260	CZ	PHE	172	4.255	75.940	-15.420	1.00	30.60
261	C	PHE	172	3.461	81.052	-13.995	1.00	33.37
262	O	PHE	172	3.910	80.861	-15.123	1.00	33.77
263	N	ARG	173	3.971	81.943	-13.161	1.00	32.63
264	CA	ARG	173	5.107	82.758	-13.555	1.00	33.71
265	CB	ARG	173	5.358	83.827	-12.493	1.00	32.59
266	CG	ARG	173	4.239	84.870	-12.425	1.00	32.04
267	CD	ARG	173	4.293	85.722	-11.159	1.00	31.29
268	NE	ARG	173	3.908	84.960	-9.973	1.00	33.00

269	CZ	ARG	173	3.859	85.465	-8.744	1.00	34.15
270	NH1	ARG	173	4.169	86.739	-8.536	1.00	34.33
271	NH2	ARG	173	3.510	84.689	-7.720	1.00	32.64
272	C	ARG	173	6.365	81.921	-13.809	1.00	35.37
273	O	ARG	173	6.437	80.748	-13.442	1.00	35.19
274	N	LEU	174	7.345	82.530	-14.464	1.00	36.50
275	CA	LEU	174	8.596	81.855	-14.784	1.00	36.82
276	CB	LEU	174	8.510	81.263	-16.190	1.00	36.87
277	CG	LEU	174	8.462	79.744	-16.387	1.00	39.47
278	CD1	LEU	174	7.723	79.069	-15.246	1.00	39.67
279	CD2	LEU	174	7.803	79.435	-17.732	1.00	36.45
280	C	LEU	174	9.725	82.879	-14.710	1.00	37.88
281	O	LEU	174	9.541	84.057	-15.033	1.00	35.22
282	N	PRO	175	10.909	82.448	-14.267	1.00	39.01
283	CD	PRO	175	11.340	81.075	-13.967	1.00	38.74
284	CA	PRO	175	12.028	83.389	-14.175	1.00	40.79
285	CB	PRO	175	13.188	82.507	-13.705	1.00	39.85
286	CG	PRO	175	12.828	81.152	-14.236	1.00	40.48
287	C	PRO	175	12.299	84.068	-15.516	1.00	43.32
288	O	PRO	175	12.268	83.425	-16.565	1.00	43.95
289	N	GLY	176	12.558	85.372	-15.466	1.00	46.21
290	CA	GLY	176	12.824	86.153	-16.666	1.00	48.00
291	C	GLY	176	13.760	85.564	-17.710	1.00	49.63
292	O	GLY	176	14.694	84.818	-17.400	1.00	49.69
293	N	VAL	177	13.493	85.919	-18.962	1.00	51.23
294	CA	VAL	177	14.280	85.462	-20.099	1.00	53.37
295	CB	VAL	177	13.374	85.151	-21.315	1.00	53.62
296	C	VAL	177	15.286	86.539	-20.503	1.00	53.91
297	O	VAL	177	16.464	86.252	-20.724	1.00	55.22
298	N	LYS	198	22.338	95.613	5.421	1.00	80.92
299	CA	LYS	198	23.394	94.734	4.927	1.00	81.15
300	CB	LYS	198	24.772	95.274	5.328	1.00	80.83
301	C	LYS	198	23.210	93.329	5.486	1.00	81.08
302	O	LYS	198	23.802	92.368	4.991	1.00	80.51
303	N	TRP	199	22.390	93.222	6.527	1.00	81.85
304	CA	TRP	199	22.111	91.936	7.151	1.00	82.25
305	CB	TRP	199	21.489	92.136	8.534	1.00	82.37
306	C	TRP	199	21.156	91.162	6.252	1.00	82.85
307	O	TRP	199	20.121	90.660	6.690	1.00	83.14
308	N	SER	200	21.513	91.102	4.975	1.00	83.20
309	CA	SER	200	20.743	90.385	3.974	1.00	82.61
310	CB	SER	200	20.490	91.281	2.758	1.00	82.34
311	C	SER	200	21.617	89.197	3.589	1.00	82.56
312	O	SER	200	21.256	88.379	2.743	1.00	82.01
313	N	GLN	201	22.783	89.122	4.224	1.00	82.96
314	CA	GLN	201	23.727	88.040	3.986	1.00	83.43
315	CB	GLN	201	25.023	88.275	4.768	1.00	82.89
316	C	GLN	201	23.066	86.756	4.457	1.00	83.63
317	O	GLN	201	23.338	85.675	3.939	1.00	83.64
318	N	VAL	202	22.201	86.884	5.458	1.00	84.04
319	CA	VAL	202	21.485	85.729	5.973	1.00	83.67
320	CB	VAL	202	21.021	85.937	7.446	1.00	83.48
321	CG1	VAL	202	19.972	87.037	7.528	1.00	82.55
322	CG2	VAL	202	20.488	84.627	8.011	1.00	81.59
323	C	VAL	202	20.283	85.570	5.054	1.00	83.89

324	O	VAL	202	19.824	84.452	4.807	1.00	84.37
325	N	ARG	203	19.799	86.700	4.531	1.00	83.35
326	CA	ARG	203	18.660	86.707	3.614	1.00	82.82
327	CB	ARG	203	18.324	88.135	3.176	1.00	82.82
328	C	ARG	203	19.060	85.873	2.408	1.00	82.34
329	O	ARG	203	18.257	85.619	1.505	1.00	81.03
330	N	LYS	204	20.323	85.462	2.410	1.00	82.10
331	CA	LYS	204	20.873	84.627	1.362	1.00	83.16
332	CB	LYS	204	22.362	84.927	1.177	1.00	82.69
333	C	LYS	204	20.671	83.178	1.798	1.00	83.94
334	O	LYS	204	21.529	82.321	1.576	1.00	83.47
335	N	ASP	205	19.528	82.924	2.435	1.00	84.86
336	CA	ASP	205	19.169	81.590	2.910	1.00	85.77
337	CB	ASP	205	17.872	81.640	3.717	1.00	85.05
338	CG	ASP	205	18.046	82.269	5.084	1.00	84.41
339	OD1	ASP	205	18.864	81.754	5.876	1.00	84.72
340	OD2	ASP	205	17.354	83.269	5.372	1.00	82.35
341	C	ASP	205	18.961	80.667	1.717	1.00	86.83
342	O	ASP	205	18.997	79.441	1.845	1.00	87.61
343	N	LEU	206	18.731	81.273	0.558	1.00	87.03
344	CA	LEU	206	18.507	80.537	-0.679	1.00	86.79
345	CB	LEU	206	17.899	81.475	-1.733	1.00	86.04
346	CG	LEU	206	17.202	82.742	-1.219	1.00	84.89
347	CD1	LEU	206	18.254	83.757	-0.806	1.00	84.34
348	CD2	LEU	206	16.316	83.339	-2.300	1.00	84.06
349	C	LEU	206	19.818	79.941	-1.203	1.00	86.85
350	O	LEU	206	19.963	79.685	-2.399	1.00	87.15
351	N	CYS	207	20.774	79.724	-0.304	1.00	86.66
352	CA	CYS	207	22.068	79.166	-0.686	1.00	86.81
353	CB	CYS	207	23.016	79.145	0.519	1.00	86.67
354	C	CYS	207	21.914	77.755	-1.246	1.00	86.37
355	O	CYS	207	22.699	77.326	-2.095	1.00	85.73
356	N	SER	208	20.904	77.038	-0.764	1.00	85.81
357	CA	SER	208	20.659	75.684	-1.236	1.00	85.03
358	CB	SER	208	20.502	74.721	-0.067	1.00	85.43
359	OG	SER	208	20.386	73.395	-0.549	1.00	86.25
360	C	SER	208	19.424	75.616	-2.120	1.00	84.59
361	O	SER	208	18.750	74.588	-2.190	1.00	84.95
362	N	LEU	209	19.118	76.730	-2.775	1.00	84.09
363	CA	LEU	209	17.993	76.796	-3.701	1.00	82.53
364	CB	LEU	209	17.426	78.222	-3.762	1.00	83.83
365	CG	LEU	209	16.071	78.490	-4.434	1.00	83.43
366	CD1	LEU	209	15.714	79.960	-4.243	1.00	84.01
367	CD2	LEU	209	16.111	78.144	-5.912	1.00	82.98
368	C	LEU	209	18.660	76.429	-5.023	1.00	80.50
369	O	LEU	209	18.687	77.212	-5.978	1.00	81.04
370	N	LYS	210	19.240	75.235	-5.045	1.00	76.77
371	CA	LYS	210	19.928	74.741	-6.223	1.00	73.31
372	CB	LYS	210	21.235	74.046	-5.816	1.00	74.78
373	CG	LYS	210	22.513	74.821	-6.153	1.00	74.99
374	CD	LYS	210	22.477	76.287	-5.688	1.00	75.47
375	CE	LYS	210	21.831	77.214	-6.728	1.00	75.10
376	NZ	LYS	210	21.959	78.662	-6.375	1.00	74.28
377	C	LYS	210	19.034	73.788	-6.989	1.00	69.16
378	O	LYS	210	19.131	72.568	-6.840	1.00	69.55

379	N	VAL	211	18.157	74.359	-7.807	1.00	63.67
380	CA	VAL	211	17.235	73.579	-8.613	1.00	58.05
381	CB	VAL	211	15.826	73.578	-8.011	1.00	57.52
382	CG1	VAL	211	15.639	72.373	-7.117	1.00	56.58
383	CG2	VAL	211	15.616	74.848	-7.230	1.00	57.72
384	C	VAL	211	17.119	74.105	-10.028	1.00	54.99
385	O	VAL	211	16.816	75.277	-10.246	1.00	53.17
386	N	SER	212	17.381	73.234	-10.994	1.00	53.23
387	CA	SER	212	17.247	73.611	-12.389	1.00	51.07
388	CB	SER	212	18.122	72.729	-13.293	1.00	50.83
389	OG	SER	212	17.552	71.451	-13.518	1.00	49.56
390	C	SER	212	15.765	73.351	-12.640	1.00	49.68
391	O	SER	212	15.130	72.598	-11.898	1.00	48.84
392	N	LEU	213	15.203	73.971	-13.663	1.00	47.72
393	CA	LEU	213	13.792	73.778	-13.927	1.00	46.62
394	CB	LEU	213	13.067	75.116	-13.774	1.00	46.11
395	CG	LEU	213	11.552	75.106	-13.580	1.00	44.82
396	CD1	LEU	213	11.104	76.412	-12.957	1.00	44.75
397	CD2	LEU	213	10.873	74.875	-14.898	1.00	45.01
398	C	LEU	213	13.591	73.210	-15.323	1.00	47.18
399	O	LEU	213	14.277	73.612	-16.262	1.00	49.18
400	N	GLN	214	12.672	72.256	-15.453	1.00	45.80
401	CA	GLN	214	12.376	71.660	-16.746	1.00	43.38
402	CB	GLN	214	12.900	70.223	-16.833	1.00	42.70
403	CG	GLN	214	12.394	69.478	-18.074	1.00	44.66
404	CD	GLN	214	13.195	68.230	-18.407	1.00	45.66
405	OE1	GLN	214	14.372	68.309	-18.749	1.00	46.09
406	NE2	GLN	214	12.554	67.071	-18.313	1.00	46.01
407	C	GLN	214	10.880	71.666	-16.994	1.00	43.78
408	O	GLN	214	10.085	71.492	-16.066	1.00	43.88
409	N	LEU	215	10.503	71.876	-18.253	1.00	43.19
410	CA	LEU	215	9.103	71.897	-18.655	1.00	42.27
411	CB	LEU	215	8.625	73.334	-18.840	1.00	41.91
412	CG	LEU	215	8.485	74.204	-17.595	1.00	41.65
413	CD1	LEU	215	9.482	75.339	-17.625	1.00	39.37
414	CD2	LEU	215	7.068	74.746	-17.552	1.00	44.25
415	C	LEU	215	8.896	71.131	-19.960	1.00	42.91
416	O	LEU	215	9.336	71.580	-21.018	1.00	40.97
417	N	ARG	216	8.231	69.978	-19.872	1.00	43.53
418	CA	ARG	216	7.946	69.147	-21.041	1.00	44.78
419	CB	ARG	216	7.631	67.705	-20.628	1.00	45.67
420	CG	ARG	216	8.792	66.725	-20.716	1.00	47.64
421	CD	ARG	216	9.533	66.637	-19.403	1.00	48.93
422	NE	ARG	216	8.668	66.159	-18.328	1.00	49.75
423	CZ	ARG	216	8.946	66.291	-17.034	1.00	49.45
424	NH1	ARG	216	10.069	66.892	-16.655	1.00	48.57
425	NH2	ARG	216	8.103	65.823	-16.122	1.00	48.81
426	C	ARG	216	6.748	69.692	-21.808	1.00	45.02
427	O	ARG	216	5.618	69.570	-21.355	1.00	45.77
428	N	GLY	217	6.989	70.282	-22.970	1.00	45.87
429	CA	GLY	217	5.885	70.811	-23.746	1.00	48.55
430	C	GLY	217	4.958	69.714	-24.241	1.00	50.78
431	O	GLY	217	5.414	68.650	-24.674	1.00	51.60
432	N	GLU	218	3.653	69.966	-24.174	1.00	52.29
433	CA	GLU	218	2.666	68.993	-24.636	1.00	54.93

434	CB	GLU	218	1.255	69.586	-24.526	1.00	53.82
435	C	GLU	218	2.966	68.596	-26.092	1.00	55.94
436	O	GLU	218	2.838	67.426	-26.477	1.00	56.03
437	N	ASP	219	3.384	69.574	-26.893	1.00	55.58
438	CA	ASP	219	3.709	69.324	-28.288	1.00	54.65
439	CB	ASP	219	3.881	70.643	-29.049	1.00	56.64
440	CG	ASP	219	5.066	71.458	-28.558	1.00	58.16
441	OD1	ASP	219	4.877	72.354	-27.705	1.00	57.15
442	OD2	ASP	219	6.193	71.194	-29.028	1.00	60.26
443	C	ASP	219	4.978	68.494	-28.423	1.00	53.48
444	O	ASP	219	5.328	68.069	-29.519	1.00	53.85
445	N	GLY	220	5.674	68.273	-27.313	1.00	52.16
446	CA	GLY	220	6.889	67.482	-27.359	1.00	50.48
447	C	GLY	220	8.157	68.267	-27.086	1.00	50.11
448	O	GLY	220	9.234	67.681	-26.893	1.00	50.52
449	N	SER	221	8.044	69.594	-27.070	1.00	47.61
450	CA	SER	221	9.214	70.430	-26.815	1.00	45.46
451	CB	SER	221	8.946	71.886	-27.227	1.00	45.64
452	OG	SER	221	7.845	72.436	-26.530	1.00	46.87
453	C	SER	221	9.608	70.369	-25.348	1.00	42.62
454	O	SER	221	8.924	69.744	-24.538	1.00	41.38
455	N	VAL	222	10.720	71.018	-25.019	1.00	41.20
456	CA	VAL	222	11.233	71.049	-23.655	1.00	38.30
457	CB	VAL	222	12.241	69.915	-23.410	1.00	36.92
458	CG1	VAL	222	12.877	70.072	-22.048	1.00	38.57
459	CG2	VAL	222	11.548	68.580	-23.501	1.00	39.37
460	C	VAL	222	11.945	72.356	-23.365	1.00	38.03
461	O	VAL	222	12.765	72.812	-24.149	1.00	38.70
462	N	TRP	223	11.617	72.970	-22.238	1.00	39.84
463	CA	TRP	223	12.271	74.206	-21.830	1.00	39.69
464	CB	TRP	223	11.242	75.287	-21.488	1.00	37.73
465	CG	TRP	223	10.793	76.136	-22.665	1.00	36.74
466	CD2	TRP	223	11.548	77.165	-23.320	1.00	34.96
467	CE2	TRP	223	10.712	77.732	-24.312	1.00	34.46
468	CE3	TRP	223	12.848	77.666	-23.163	1.00	34.92
469	CD1	TRP	223	9.565	76.119	-23.278	1.00	36.20
470	NE1	TRP	223	9.511	77.076	-24.263	1.00	33.37
471	CZ2	TRP	223	11.137	78.777	-25.145	1.00	34.03
472	CZ3	TRP	223	13.270	78.711	-23.994	1.00	34.63
473	CH2	TRP	223	12.414	79.252	-24.970	1.00	32.85
474	C	TRP	223	13.077	73.838	-20.590	1.00	40.80
475	O	TRP	223	12.570	73.166	-19.693	1.00	41.14
476	N	ASN	224	14.332	74.258	-20.544	1.00	42.30
477	CA	ASN	224	15.177	73.944	-19.405	1.00	42.96
478	CB	ASN	224	16.275	72.974	-19.827	1.00	44.51
479	CG	ASN	224	16.324	71.737	-18.954	1.00	45.45
480	OD1	ASN	224	16.724	71.799	-17.790	1.00	46.64
481	ND2	ASN	224	15.908	70.602	-19.512	1.00	44.92
482	C	ASN	224	15.790	75.213	-18.856	1.00	44.20
483	O	ASN	224	16.420	75.979	-19.592	1.00	44.36
484	N	TYR	225	15.601	75.425	-17.557	1.00	44.67
485	CA	TYR	225	16.113	76.609	-16.884	1.00	45.33
486	CB	TYR	225	14.974	77.360	-16.201	1.00	44.44
487	CG	TYR	225	15.446	78.556	-15.406	1.00	44.12
488	CD1	TYR	225	15.823	79.730	-16.042	1.00	42.37

489	CE1	TYR	225	16.282	80.824	-15.321	1.00	43.63
490	CD2	TYR	225	15.539	78.501	-14.015	1.00	45.39
491	CE2	TYR	225	16.002	79.597	-13.279	1.00	45.19
492	CZ	TYR	225	16.372	80.752	-13.942	1.00	44.84
493	OH	TYR	225	16.855	81.830	-13.237	1.00	45.83
494	C	TYR	225	17.182	76.313	-15.844	1.00	47.27
495	O	TYR	225	16.969	75.525	-14.922	1.00	47.12
496	N	LYS	226	18.331	76.962	-16.000	1.00	49.48
497	CA	LYS	226	19.439	76.818	-15.068	1.00	51.21
498	CB	LYS	226	20.707	76.408	-15.819	1.00	51.78
499	CG	LYS	226	21.927	76.189	-14.935	1.00	54.89
500	CD	LYS	226	22.686	77.492	-14.689	1.00	55.99
501	CE	LYS	226	23.924	77.264	-13.838	1.00	57.16
502	NZ	LYS	226	23.574	76.774	-12.473	1.00	58.76
503	C	LYS	226	19.604	78.188	-14.408	1.00	51.63
504	O	LYS	226	19.716	79.204	-15.093	1.00	50.53
505	N	PRO	227	19.604	78.230	-13.065	1.00	52.69
506	CD	PRO	227	19.452	77.066	-12.174	1.00	52.75
507	CA	PRO	227	19.743	79.465	-12.283	1.00	53.75
508	CB	PRO	227	19.260	79.045	-10.904	1.00	52.34
509	CG	PRO	227	19.774	77.654	-10.812	1.00	53.76
510	C	PRO	227	21.142	80.075	-12.239	1.00	55.58
511	O	PRO	227	22.143	79.398	-12.461	1.00	54.41
512	N	PRO	228	21.219	81.377	-11.934	1.00	58.75
513	CD	PRO	228	20.054	82.251	-11.698	1.00	60.66
514	CA	PRO	228	22.468	82.139	-11.844	1.00	62.36
515	CB	PRO	228	21.976	83.583	-11.777	1.00	63.06
516	CG	PRO	228	20.675	83.458	-11.034	1.00	61.42
517	C	PRO	228	23.345	81.772	-10.650	1.00	65.20
518	O	PRO	228	23.437	80.608	-10.268	1.00	65.74
519	N	ALA	229	23.993	82.782	-10.075	1.00	69.06
520	CA	ALA	229	24.867	82.601	-8.917	1.00	71.82
521	CB	ALA	229	26.318	82.470	-9.371	1.00	70.23
522	C	ALA	229	24.720	83.789	-7.964	1.00	74.46
523	O	ALA	229	24.219	84.850	-8.349	1.00	74.74
524	N	ASP	230	25.152	83.605	-6.720	1.00	77.21
525	CA	ASP	230	25.068	84.666	-5.725	1.00	79.67
526	CB	ASP	230	25.096	84.070	-4.314	1.00	78.65
527	C	ASP	230	26.231	85.636	-5.917	1.00	82.10
528	O	ASP	230	26.632	85.909	-7.054	1.00	81.03
529	N	SER	231	26.772	86.147	-4.809	1.00	85.79
530	CA	SER	231	27.887	87.105	-4.835	1.00	87.83
531	CB	SER	231	29.188	86.413	-5.261	1.00	88.27
532	C	SER	231	27.526	88.226	-5.806	1.00	88.95
533	O	SER	231	28.340	89.091	-6.147	1.00	88.98
534	N	GLY	232	26.276	88.157	-6.245	1.00	89.33
535	CA	GLY	232	25.674	89.117	-7.137	1.00	90.02
536	C	GLY	232	24.354	89.188	-6.409	1.00	90.80
537	O	GLY	232	23.279	89.024	-6.987	1.00	91.71
538	N	GLY	233	24.477	89.383	-5.096	1.00	90.52
539	CA	GLY	233	23.338	89.459	-4.197	1.00	89.63
540	C	GLY	233	22.073	90.063	-4.766	1.00	88.36
541	O	GLY	233	20.970	89.692	-4.361	1.00	89.14
542	N	LYS	234	22.220	90.995	-5.699	1.00	86.18
543	CA	LYS	234	21.062	91.632	-6.303	1.00	83.99

544	CB	LYS	234	21.492	92.884	-7.075	1.00	83.43
545	C	LYS	234	20.322	90.676	-7.236	1.00	82.47
546	O	LYS	234	20.864	90.243	-8.251	1.00	83.25
547	N	GLU	235	19.090	90.340	-6.871	1.00	80.37
548	CA	GLU	235	18.244	89.461	-7.676	1.00	78.49
549	CB	GLU	235	18.052	90.061	-9.072	1.00	79.29
550	C	GLU	235	18.669	87.998	-7.814	1.00	77.17
551	O	GLU	235	19.144	87.562	-8.866	1.00	76.76
552	N	ILE	236	18.490	87.248	-6.736	1.00	75.47
553	CA	ILE	236	18.781	85.819	-6.710	1.00	72.63
554	CB	ILE	236	19.778	85.461	-5.586	1.00	73.98
555	CG2	ILE	236	21.206	85.691	-6.065	1.00	74.10
556	CG1	ILE	236	19.458	86.280	-4.328	1.00	74.81
557	CD1	ILE	236	20.463	86.106	-3.189	1.00	75.72
558	C	ILE	236	17.414	85.228	-6.399	1.00	69.92
559	O	ILE	236	17.280	84.069	-5.997	1.00	69.58
560	N	PHE	237	16.406	86.077	-6.587	1.00	66.10
561	CA	PHE	237	15.010	85.739	-6.356	1.00	61.51
562	CB	PHE	237	14.291	86.898	-5.667	1.00	62.12
563	CG	PHE	237	15.115	87.589	-4.628	1.00	60.86
564	CD1	PHE	237	15.850	88.723	-4.951	1.00	59.98
565	CD2	PHE	237	15.177	87.089	-3.331	1.00	59.56
566	CE1	PHE	237	16.637	89.348	-3.997	1.00	60.01
567	CE2	PHE	237	15.957	87.702	-2.373	1.00	59.37
568	CZ	PHE	237	16.691	88.836	-2.704	1.00	60.00
569	C	PHE	237	14.348	85.491	-7.696	1.00	58.40
570	O	PHE	237	13.124	85.511	-7.812	1.00	58.37
571	N	SER	238	15.169	85.279	-8.716	1.00	54.37
572	CA	SER	238	14.657	85.028	-10.049	1.00	49.97
573	CB	SER	238	15.824	84.698	-10.990	1.00	50.06
574	OG	SER	238	16.776	83.867	-10.349	1.00	50.94
575	C	SER	238	13.627	83.896	-10.042	1.00	46.05
576	O	SER	238	12.686	83.896	-10.837	1.00	44.82
577	N	LEU	239	13.782	82.956	-9.115	1.00	41.74
578	CA	LEU	239	12.877	81.819	-9.037	1.00	37.63
579	CB	LEU	239	13.703	80.554	-8.827	1.00	38.30
580	CG	LEU	239	13.323	79.289	-9.603	1.00	40.67
581	CD1	LEU	239	12.910	79.620	-11.030	1.00	40.00
582	CD2	LEU	239	14.520	78.346	-9.605	1.00	41.71
583	C	LEU	239	11.759	81.893	-7.986	1.00	35.73
584	O	LEU	239	10.905	81.017	-7.949	1.00	35.63
585	N	LEU	240	11.743	82.929	-7.150	1.00	33.56
586	CA	LEU	240	10.713	83.057	-6.110	1.00	31.58
587	CB	LEU	240	10.948	84.294	-5.229	1.00	32.56
588	CG	LEU	240	11.966	84.203	-4.095	1.00	32.81
589	CD1	LEU	240	11.841	85.443	-3.232	1.00	34.47
590	CD2	LEU	240	11.715	82.961	-3.262	1.00	33.77
591	C	LEU	240	9.280	83.117	-6.594	1.00	29.50
592	O	LEU	240	8.439	82.349	-6.144	1.00	30.03
593	N	PRO	241	8.970	84.052	-7.500	1.00	28.88
594	CD	PRO	241	9.866	84.966	-8.230	1.00	27.35
595	CA	PRO	241	7.593	84.154	-7.993	1.00	29.16
596	CB	PRO	241	7.712	85.125	-9.166	1.00	27.20
597	CG	PRO	241	8.902	85.951	-8.813	1.00	26.11
598	C	PRO	241	7.024	82.800	-8.417	1.00	31.71

599	O	PRO	241	5.997	82.360	-7.892	1.00	32.87
600	N	HIS	242	7.690	82.153	-9.374	1.00	32.44
601	CA	HIS	242	7.273	80.840	-9.869	1.00	33.09
602	CB	HIS	242	8.376	80.183	-10.696	1.00	33.12
603	CG	HIS	242	8.151	78.722	-10.931	1.00	31.81
604	CD2	HIS	242	8.575	77.633	-10.246	1.00	33.86
605	ND1	HIS	242	7.315	78.251	-11.916	1.00	33.72
606	CE1	HIS	242	7.228	76.935	-11.828	1.00	34.99
607	NE2	HIS	242	7.982	76.535	-10.821	1.00	34.87
608	C	HIS	242	7.010	79.923	-8.705	1.00	32.78
609	O	HIS	242	5.955	79.300	-8.603	1.00	33.32
610	N	MET	243	8.027	79.830	-7.859	1.00	34.86
611	CA	MET	243	8.016	78.989	-6.681	1.00	37.25
612	CB	MET	243	9.270	79.237	-5.843	1.00	41.17
613	CG	MET	243	10.541	78.739	-6.445	1.00	48.42
614	SD	MET	243	10.419	76.969	-6.614	1.00	56.97
615	CE	MET	243	11.078	76.678	-8.280	1.00	58.78
616	C	MET	243	6.785	79.266	-5.866	1.00	36.91
617	O	MET	243	6.110	78.336	-5.445	1.00	36.25
618	N	ALA	244	6.505	80.541	-5.615	1.00	36.65
619	CA	ALA	244	5.316	80.924	-4.846	1.00	37.41
620	CB	ALA	244	5.249	82.457	-4.683	1.00	37.71
621	C	ALA	244	4.017	80.424	-5.490	1.00	36.31
622	O	ALA	244	3.086	80.012	-4.799	1.00	37.79
623	N	ASP	245	3.940	80.475	-6.814	1.00	35.19
624	CA	ASP	245	2.742	80.016	-7.497	1.00	34.32
625	CB	ASP	245	2.810	80.355	-8.985	1.00	31.56
626	CG	ASP	245	2.738	81.849	-9.235	1.00	30.74
627	OD1	ASP	245	2.014	82.528	-8.482	1.00	28.28
628	OD2	ASP	245	3.389	82.346	-10.180	1.00	30.84
629	C	ASP	245	2.529	78.525	-7.295	1.00	34.63
630	O	ASP	245	1.390	78.085	-7.105	1.00	36.05
631	N	MET	246	3.619	77.760	-7.323	1.00	34.62
632	CA	MET	246	3.571	76.306	-7.126	1.00	35.13
633	CB	MET	246	4.946	75.686	-7.378	1.00	35.05
634	CG	MET	246	5.028	74.211	-7.042	1.00	35.60
635	SD	MET	246	4.127	73.170	-8.205	1.00	36.53
636	CE	MET	246	2.599	73.067	-7.434	1.00	37.78
637	C	MET	246	3.139	75.989	-5.697	1.00	34.29
638	O	MET	246	2.446	75.006	-5.436	1.00	33.03
639	N	SER	247	3.566	76.835	-4.774	1.00	33.90
640	CA	SER	247	3.215	76.663	-3.378	1.00	34.79
641	CB	SER	247	3.991	77.651	-2.525	1.00	37.25
642	OG	SER	247	5.349	77.302	-2.533	1.00	41.24
643	C	SER	247	1.737	76.897	-3.171	1.00	32.81
644	O	SER	247	1.084	76.218	-2.380	1.00	32.97
645	N	THR	248	1.214	77.881	-3.882	1.00	32.23
646	CA	THR	248	-0.190	78.218	-3.772	1.00	31.48
647	CB	THR	248	-0.494	79.492	-4.549	1.00	32.12
648	OG1	THR	248	0.321	80.553	-4.035	1.00	33.82
649	CG2	THR	248	-1.961	79.864	-4.410	1.00	31.31
650	C	THR	248	-1.026	77.075	-4.310	1.00	30.39
651	O	THR	248	-2.012	76.677	-3.693	1.00	29.27
652	N	TYR	249	-0.622	76.546	-5.462	1.00	30.38
653	CA	TYR	249	-1.329	75.434	-6.076	1.00	30.33

654	CB	TYR	249	-0.649	75.034	-7.386	1.00	29.81
655	CG	TYR	249	-1.224	73.784	-8.008	1.00	30.36
656	CD1	TYR	249	-2.531	73.766	-8.498	1.00	28.62
657	CE1	TYR	249	-3.086	72.601	-8.996	1.00	31.63
658	CD2	TYR	249	-0.482	72.594	-8.043	1.00	28.99
659	CE2	TYR	249	-1.018	71.424	-8.533	1.00	29.89
660	CZ	TYR	249	-2.328	71.426	-9.008	1.00	34.44
661	OH	TYR	249	-2.891	70.247	-9.459	1.00	36.73
662	C	TYR	249	-1.324	74.270	-5.086	1.00	31.21
663	O	TYR	249	-2.360	73.642	-4.836	1.00	30.19
664	N	MET	250	-0.155	73.996	-4.510	1.00	31.98
665	CA	MET	250	-0.028	72.921	-3.527	1.00	33.11
666	CB	MET	250	1.432	72.750	-3.110	1.00	32.37
667	CG	MET	250	2.291	72.132	-4.186	1.00	31.75
668	SD	MET	250	1.684	70.502	-4.687	1.00	34.03
669	CE	MET	250	2.509	69.452	-3.449	1.00	31.25
670	C	MET	250	-0.885	73.188	-2.291	1.00	33.11
671	O	MET	250	-1.629	72.323	-1.836	1.00	32.66
672	N	PHE	251	-0.775	74.390	-1.745	1.00	33.37
673	CA	PHE	251	-1.561	74.744	-0.576	1.00	33.59
674	CB	PHE	251	-1.286	76.195	-0.188	1.00	34.12
675	CG	PHE	251	0.077	76.411	0.392	1.00	34.59
676	CD1	PHE	251	0.700	77.647	0.287	1.00	34.63
677	CD2	PHE	251	0.724	75.385	1.076	1.00	35.08
678	CE1	PHE	251	1.949	77.863	0.855	1.00	36.50
679	CE2	PHE	251	1.972	75.587	1.648	1.00	35.30
680	CZ	PHE	251	2.585	76.834	1.537	1.00	36.25
681	C	PHE	251	-3.044	74.540	-0.864	1.00	33.35
682	O	PHE	251	-3.742	73.869	-0.102	1.00	32.53
683	N	LYS	252	-3.522	75.117	-1.964	1.00	32.74
684	CA	LYS	252	-4.921	74.972	-2.337	1.00	31.51
685	CB	LYS	252	-5.136	75.467	-3.766	1.00	31.30
686	CG	LYS	252	-5.214	76.978	-3.879	1.00	33.17
687	CD	LYS	252	-5.138	77.454	-5.324	1.00	36.80
688	CE	LYS	252	-5.911	78.769	-5.533	1.00	40.52
689	NZ	LYS	252	-5.647	79.815	-4.486	1.00	42.52
690	C	LYS	252	-5.318	73.503	-2.214	1.00	32.28
691	O	LYS	252	-6.335	73.174	-1.606	1.00	32.02
692	N	LYS	253	-4.493	72.622	-2.774	1.00	32.25
693	CA	LYS	253	-4.750	71.190	-2.725	1.00	31.71
694	C	LYS	253	-4.820	70.699	-1.281	1.00	33.18
695	O	LYS	253	-5.708	69.911	-0.926	1.00	34.96
696	N	ILE	254	-3.880	71.164	-0.458	1.00	32.64
697	CA	ILE	254	-3.807	70.797	0.960	1.00	31.44
698	CB	ILE	254	-2.495	71.361	1.609	1.00	31.24
699	CG2	ILE	254	-2.757	71.843	3.025	1.00	30.88
700	CG1	ILE	254	-1.425	70.273	1.667	1.00	31.57
701	CD1	ILE	254	-1.108	69.641	0.345	1.00	35.02
702	C	ILE	254	-5.030	71.272	1.772	1.00	30.82
703	O	ILE	254	-5.500	70.568	2.671	1.00	30.56
704	N	ILE	255	-5.533	72.465	1.471	1.00	29.36
705	CA	ILE	255	-6.696	72.982	2.190	1.00	29.54
706	CB	ILE	255	-6.970	74.485	1.850	1.00	29.42
707	CG2	ILE	255	-8.291	74.939	2.449	1.00	24.21
708	CG1	ILE	255	-5.829	75.357	2.383	1.00	29.56

709	CD1	ILE	255	-5.954	76.826	1.993	1.00	29.49
710	C	ILE	255	-7.902	72.140	1.798	1.00	28.95
711	O	ILE	255	-8.761	71.841	2.626	1.00	28.74
712	N	SER	256	-7.952	71.752	0.529	1.00	28.49
713	CA	SER	256	-9.046	70.935	0.040	1.00	29.78
714	CB	SER	256	-8.946	70.765	-1.473	1.00	29.69
715	OG	SER	256	-8.929	72.020	-2.116	1.00	31.07
716	C	SER	256	-8.992	69.569	0.714	1.00	29.53
717	O	SER	256	-10.015	68.930	0.924	1.00	29.42
718	N	PHE	257	-7.790	69.119	1.042	1.00	29.28
719	CA	PHE	257	-7.628	67.832	1.699	1.00	30.50
720	CB	PHE	257	-6.145	67.504	1.855	1.00	29.19
721	CG	PHE	257	-5.887	66.239	2.611	1.00	28.46
722	CD1	PHE	257	-6.234	65.001	2.067	1.00	28.14
723	CD2	PHE	257	-5.309	66.281	3.876	1.00	28.50
724	CE1	PHE	257	-6.008	63.814	2.776	1.00	27.99
725	CE2	PHE	257	-5.078	65.102	4.597	1.00	28.73
726	CZ	PHE	257	-5.430	63.866	4.042	1.00	29.76
727	C	PHE	257	-8.269	67.901	3.075	1.00	31.48
728	O	PHE	257	-9.085	67.054	3.450	1.00	29.83
729	N	ALA	258	-7.880	68.924	3.824	1.00	33.28
730	CA	ALA	258	-8.387	69.137	5.167	1.00	34.73
731	CB	ALA	258	-7.792	70.418	5.743	1.00	33.79
732	C	ALA	258	-9.910	69.208	5.193	1.00	36.37
733	O	ALA	258	-10.565	68.399	5.843	1.00	36.08
734	N	LYS	259	-10.464	70.175	4.470	1.00	38.81
735	CA	LYS	259	-11.908	70.393	4.420	1.00	40.37
736	CB	LYS	259	-12.208	71.614	3.553	1.00	39.71
737	CG	LYS	259	-11.669	72.926	4.093	1.00	41.17
738	CD	LYS	259	-11.994	74.064	3.135	1.00	43.90
739	CE	LYS	259	-11.760	75.433	3.769	1.00	48.08
740	NZ	LYS	259	-12.688	75.721	4.916	1.00	49.91
741	C	LYS	259	-12.742	69.217	3.912	1.00	42.23
742	O	LYS	259	-13.970	69.236	4.027	1.00	42.77
743	N	VAL	260	-12.092	68.196	3.359	1.00	42.41
744	CA	VAL	260	-12.820	67.056	2.817	1.00	43.19
745	CB	VAL	260	-12.247	66.621	1.450	1.00	44.42
746	CG1	VAL	260	-10.911	65.914	1.646	1.00	45.71
747	CG2	VAL	260	-13.228	65.722	0.733	1.00	44.46
748	C	VAL	260	-12.813	65.846	3.730	1.00	44.07
749	O	VAL	260	-13.396	64.812	3.393	1.00	43.88
750	N	ILE	261	-12.140	65.958	4.873	1.00	44.41
751	CA	ILE	261	-12.101	64.848	5.817	1.00	44.20
752	CB	ILE	261	-10.928	64.998	6.838	1.00	43.58
753	CG2	ILE	261	-11.054	63.971	7.953	1.00	43.03
754	CG1	ILE	261	-9.587	64.796	6.107	1.00	44.73
755	CD1	ILE	261	-8.362	64.682	7.006	1.00	44.15
756	C	ILE	261	-13.461	64.799	6.514	1.00	44.53
757	O	ILE	261	-14.441	64.361	5.910	1.00	47.51
758	N	SER	262	-13.536	65.246	7.759	1.00	42.88
759	CA	SER	262	-14.794	65.257	8.507	1.00	42.04
760	CB	SER	262	-15.533	63.915	8.408	1.00	39.59
761	OG	SER	262	-14.948	62.919	9.222	1.00	39.47
762	C	SER	262	-14.420	65.545	9.950	1.00	42.75
763	O	SER	262	-14.984	66.434	10.593	1.00	42.22

764	N	TYR	263	-13.455	64.797	10.464	1.00	42.42
765	CA	TYR	263	-13.015	65.053	11.811	1.00	43.67
766	CB	TYR	263	-11.820	64.174	12.166	1.00	45.14
767	CG	TYR	263	-12.052	62.695	11.952	1.00	47.82
768	CD1	TYR	263	-11.991	62.136	10.674	1.00	47.67
769	CE1	TYR	263	-12.184	60.773	10.473	1.00	50.18
770	CD2	TYR	263	-12.319	61.850	13.028	1.00	48.79
771	CE2	TYR	263	-12.516	60.481	12.839	1.00	50.99
772	CZ	TYR	263	-12.446	59.948	11.559	1.00	51.85
773	OH	TYR	263	-12.631	58.592	11.368	1.00	52.57
774	C	TYR	263	-12.597	66.522	11.796	1.00	43.98
775	O	TYR	263	-12.695	67.214	12.805	1.00	45.50
776	N	PHE	264	-12.160	67.002	10.631	1.00	43.31
777	CA	PHE	264	-11.717	68.388	10.490	1.00	41.84
778	CB	PHE	264	-10.842	68.542	9.238	1.00	40.15
779	CG	PHE	264	-10.114	69.857	9.165	1.00	38.19
780	CD1	PHE	264	-9.078	70.144	10.047	1.00	38.14
781	CD2	PHE	264	-10.480	70.822	8.228	1.00	37.78
782	CE1	PHE	264	-8.414	71.382	9.999	1.00	38.64
783	CE2	PHE	264	-9.823	72.063	8.170	1.00	37.53
784	CZ	PHE	264	-8.791	72.342	9.056	1.00	37.70
785	C	PHE	264	-12.881	69.379	10.433	1.00	41.17
786	O	PHE	264	-12.877	70.383	11.133	1.00	38.98
787	N	ARG	265	-13.876	69.090	9.603	1.00	42.13
788	CA	ARG	265	-15.029	69.971	9.470	1.00	44.16
789	CB	ARG	265	-15.956	69.488	8.352	1.00	47.70
790	CG	ARG	265	-15.410	69.686	6.945	1.00	52.95
791	CD	ARG	265	-16.506	69.516	5.901	1.00	56.48
792	NE	ARG	265	-17.185	68.224	6.007	1.00	60.21
793	CZ	ARG	265	-16.625	67.050	5.730	1.00	61.99
794	NH1	ARG	265	-15.361	66.996	5.323	1.00	63.67
795	NH2	ARG	265	-17.329	65.927	5.858	1.00	61.76
796	C	ARG	265	-15.842	70.112	10.742	1.00	44.25
797	O	ARG	265	-16.562	71.096	10.908	1.00	44.19
798	N	ASP	266	-15.750	69.127	11.633	1.00	44.83
799	CA	ASP	266	-16.493	69.189	12.889	1.00	44.68
800	CB	ASP	266	-16.495	67.827	13.603	1.00	45.71
801	CG	ASP	266	-17.304	66.758	12.855	1.00	49.06
802	OD1	ASP	266	-18.266	67.112	12.131	1.00	48.27
803	OD2	ASP	266	-16.987	65.551	13.006	1.00	49.66
804	C	ASP	266	-15.880	70.249	13.805	1.00	44.33
805	O	ASP	266	-16.576	70.860	14.620	1.00	45.11
806	N	LEU	267	-14.577	70.472	13.663	1.00	43.41
807	CA	LEU	267	-13.877	71.461	14.474	1.00	43.59
808	CB	LEU	267	-12.369	71.417	14.188	1.00	42.43
809	CG	LEU	267	-11.613	70.082	14.287	1.00	41.73
810	CD1	LEU	267	-10.204	70.272	13.767	1.00	39.48
811	CD2	LEU	267	-11.580	69.578	15.727	1.00	39.34
812	C	LEU	267	-14.409	72.852	14.131	1.00	44.86
813	O	LEU	267	-14.834	73.102	12.999	1.00	44.76
814	N	PRO	268	-14.416	73.774	15.106	1.00	44.92
815	CD	PRO	268	-14.049	73.664	16.528	1.00	45.28
816	CA	PRO	268	-14.911	75.115	14.798	1.00	45.77
817	CB	PRO	268	-14.908	75.806	16.160	1.00	45.23
818	CG	PRO	268	-13.804	75.107	16.897	1.00	44.19

819	C	PRO	268	-14.004	75.810	13.775	1.00	47.17
820	O	PRO	268	-12.789	75.589	13.751	1.00	46.04
821	N	ILE	269	-14.608	76.645	12.937	1.00	48.42
822	CA	ILE	269	-13.890	77.372	11.898	1.00	50.36
823	CB	ILE	269	-14.680	78.609	11.414	1.00	51.28
824	CG2	ILE	269	-14.150	79.038	10.058	1.00	51.69
825	CG1	ILE	269	-16.180	78.295	11.303	1.00	53.81
826	CD1	ILE	269	-16.955	78.345	12.638	1.00	54.19
827	C	ILE	269	-12.497	77.857	12.301	1.00	50.55
828	O	ILE	269	-11.500	77.483	11.686	1.00	49.86
829	N	GLU	270	-12.429	78.694	13.330	1.00	51.84
830	CA	GLU	270	-11.153	79.241	13.783	1.00	52.67
831	CB	GLU	270	-11.340	80.074	15.057	1.00	55.61
832	CG	GLU	270	-12.428	81.130	14.949	1.00	59.36
833	CD	GLU	270	-13.795	80.604	15.353	1.00	62.33
834	OE1	GLU	270	-14.166	79.488	14.920	1.00	63.26
835	OE2	GLU	270	-14.501	81.315	16.102	1.00	64.35
836	C	GLU	270	-10.110	78.163	14.030	1.00	51.41
837	O	GLU	270	-8.917	78.383	13.805	1.00	50.49
838	N	ASP	271	-10.559	77.000	14.497	1.00	50.41
839	CA	ASP	271	-9.645	75.898	14.761	1.00	49.31
840	CB	ASP	271	-10.327	74.825	15.613	1.00	50.84
841	CG	ASP	271	-10.098	75.042	17.093	1.00	52.85
842	OD1	ASP	271	-10.538	74.206	17.909	1.00	55.88
843	OD2	ASP	271	-9.463	76.059	17.441	1.00	54.72
844	C	ASP	271	-9.119	75.299	13.467	1.00	47.56
845	O	ASP	271	-8.004	74.781	13.419	1.00	47.18
846	N	GLN	272	-9.926	75.374	12.415	1.00	45.59
847	CA	GLN	272	-9.512	74.862	11.119	1.00	43.42
848	CB	GLN	272	-10.687	74.846	10.148	1.00	42.45
849	CG	GLN	272	-11.818	73.967	10.613	1.00	41.49
850	CD	GLN	272	-12.943	73.912	9.619	1.00	41.93
851	OE1	GLN	272	-13.899	73.163	9.796	1.00	43.50
852	NE2	GLN	272	-12.839	74.706	8.560	1.00	44.63
853	C	GLN	272	-8.425	75.783	10.595	1.00	41.78
854	O	GLN	272	-7.375	75.322	10.155	1.00	42.19
855	N	ILE	273	-8.689	77.085	10.652	1.00	39.31
856	CA	ILE	273	-7.732	78.085	10.201	1.00	38.30
857	CB	ILE	273	-8.255	79.529	10.461	1.00	37.96
858	CG2	ILE	273	-7.177	80.551	10.135	1.00	34.75
859	CG1	ILE	273	-9.526	79.790	9.649	1.00	37.60
860	CD1	ILE	273	-9.377	79.509	8.168	1.00	39.57
861	C	ILE	273	-6.422	77.903	10.967	1.00	39.07
862	O	ILE	273	-5.333	77.881	10.385	1.00	37.83
863	N	SER	274	-6.523	77.763	12.281	1.00	38.53
864	CA	SER	274	-5.317	77.612	13.070	1.00	38.45
865	CB	SER	274	-5.646	77.571	14.562	1.00	39.70
866	OG	SER	274	-6.094	78.842	15.000	1.00	41.30
867	C	SER	274	-4.530	76.385	12.667	1.00	36.53
868	O	SER	274	-3.346	76.490	12.370	1.00	38.41
869	N	LEU	275	-5.180	75.227	12.639	1.00	33.73
870	CA	LEU	275	-4.480	73.999	12.274	1.00	32.28
871	CB	LEU	275	-5.429	72.795	12.377	1.00	29.43
872	CG	LEU	275	-5.954	72.520	13.799	1.00	26.80
873	CD1	LEU	275	-7.045	71.477	13.749	1.00	22.87

874	CD2	LEU	275	-4.822	72.078	14.706	1.00	22.58
875	C	LEU	275	-3.864	74.092	10.876	1.00	32.13
876	O	LEU	275	-2.739	73.619	10.654	1.00	31.06
877	N	LEU	276	-4.581	74.732	9.951	1.00	30.82
878	CA	LEU	276	-4.099	74.892	8.583	1.00	31.04
879	CB	LEU	276	-5.180	75.507	7.687	1.00	28.88
880	CG	LEU	276	-6.315	74.554	7.281	1.00	31.13
881	CD1	LEU	276	-7.333	75.264	6.378	1.00	29.42
882	CD2	LEU	276	-5.713	73.336	6.564	1.00	28.71
883	C	LEU	276	-2.838	75.736	8.507	1.00	32.66
884	O	LEU	276	-1.854	75.332	7.886	1.00	33.47
885	N	LYS	277	-2.859	76.905	9.139	1.00	34.74
886	CA	LYS	277	-1.694	77.786	9.117	1.00	34.78
887	CB	LYS	277	-1.980	79.094	9.869	1.00	35.39
888	CG	LYS	277	-3.277	79.769	9.432	1.00	38.47
889	CD	LYS	277	-3.200	81.286	9.449	1.00	39.03
890	CE	LYS	277	-3.118	81.846	10.853	1.00	41.48
891	NZ	LYS	277	-3.043	83.342	10.813	1.00	41.41
892	C	LYS	277	-0.542	77.050	9.773	1.00	35.62
893	O	LYS	277	0.627	77.280	9.447	1.00	38.27
894	N	GLY	278	-0.874	76.143	10.683	1.00	33.80
895	CA	GLY	278	0.159	75.398	11.366	1.00	34.06
896	C	GLY	278	0.758	74.245	10.578	1.00	35.11
897	O	GLY	278	1.952	73.963	10.702	1.00	35.77
898	N	ALA	279	-0.043	73.588	9.744	1.00	34.64
899	CA	ALA	279	0.460	72.432	9.002	1.00	32.85
900	CB	ALA	279	-0.448	71.216	9.288	1.00	30.51
901	C	ALA	279	0.678	72.573	7.491	1.00	30.96
902	O	ALA	279	1.276	71.690	6.873	1.00	31.54
903	N	ALA	280	0.205	73.663	6.897	1.00	29.10
904	CA	ALA	280	0.352	73.873	5.455	1.00	28.42
905	CB	ALA	280	0.140	75.346	5.122	1.00	27.99
906	C	ALA	280	1.696	73.394	4.890	1.00	28.34
907	O	ALA	280	1.737	72.498	4.045	1.00	29.63
908	N	PHE	281	2.790	73.983	5.363	1.00	27.68
909	CA	PHE	281	4.128	73.620	4.897	1.00	28.10
910	CB	PHE	281	5.189	74.320	5.753	1.00	27.65
911	CG	PHE	281	6.608	73.942	5.406	1.00	26.90
912	CD1	PHE	281	7.225	74.457	4.274	1.00	28.51
913	CD2	PHE	281	7.329	73.061	6.203	1.00	28.63
914	CE1	PHE	281	8.544	74.110	3.924	1.00	27.70
915	CE2	PHE	281	8.660	72.700	5.862	1.00	30.96
916	CZ	PHE	281	9.262	73.233	4.716	1.00	27.12
917	C	PHE	281	4.386	72.113	4.933	1.00	29.79
918	O	PHE	281	4.815	71.518	3.932	1.00	28.38
919	N	GLU	282	4.136	71.512	6.096	1.00	28.96
920	CA	GLU	282	4.357	70.084	6.315	1.00	30.28
921	CB	GLU	282	4.037	69.737	7.772	1.00	31.58
922	CG	GLU	282	5.109	70.204	8.750	1.00	33.01
923	CD	GLU	282	4.605	70.333	10.173	1.00	33.56
924	OE1	GLU	282	4.053	69.350	10.712	1.00	33.93
925	OE2	GLU	282	4.770	71.424	10.757	1.00	33.18
926	C	GLU	282	3.571	69.171	5.380	1.00	30.84
927	O	GLU	282	4.120	68.238	4.796	1.00	27.84
928	N	LEU	283	2.279	69.446	5.252	1.00	33.57

929	CA	LEU	283	1.406	68.664	4.394	1.00	34.52
930	CB	LEU	283	-0.018	69.166	4.550	1.00	36.39
931	CG	LEU	283	-0.563	69.003	5.968	1.00	36.84
932	CD1	LEU	283	-1.903	69.723	6.063	1.00	38.02
933	CD2	LEU	283	-0.695	67.514	6.307	1.00	34.04
934	C	LEU	283	1.839	68.760	2.937	1.00	35.27
935	O	LEU	283	1.785	67.777	2.196	1.00	34.71
936	N	CYS	284	2.269	69.949	2.531	1.00	35.88
937	CA	CYS	284	2.722	70.166	1.160	1.00	36.95
938	CB	CYS	284	2.771	71.665	0.870	1.00	37.59
939	SG	CYS	284	3.645	72.045	-0.657	1.00	47.17
940	C	CYS	284	4.106	69.540	0.942	1.00	35.43
941	O	CYS	284	4.435	69.038	-0.139	1.00	33.42
942	N	GLN	285	4.899	69.580	2.005	1.00	35.36
943	CA	GLN	285	6.246	69.042	2.029	1.00	33.83
944	CB	GLN	285	6.897	69.408	3.352	1.00	36.11
945	CG	GLN	285	8.370	69.542	3.241	1.00	38.05
946	CD	GLN	285	8.757	70.365	2.038	1.00	38.90
947	OE1	GLN	285	9.363	69.863	1.083	1.00	39.51
948	NE2	GLN	285	8.400	71.637	2.067	1.00	39.43
949	C	GLN	285	6.177	67.531	1.882	1.00	33.35
950	O	GLN	285	6.955	66.931	1.145	1.00	34.22
951	N	LEU	286	5.223	66.932	2.591	1.00	33.14
952	CA	LEU	286	4.985	65.490	2.570	1.00	30.58
953	CB	LEU	286	3.937	65.122	3.619	1.00	29.89
954	CG	LEU	286	4.358	65.292	5.077	1.00	30.57
955	CD1	LEU	286	3.145	65.167	5.996	1.00	27.00
956	CD2	LEU	286	5.415	64.247	5.408	1.00	29.51
957	C	LEU	286	4.495	65.017	1.208	1.00	30.01
958	O	LEU	286	4.860	63.935	0.752	1.00	29.71
959	N	ARG	287	3.653	65.820	0.564	1.00	28.87
960	CA	ARG	287	3.131	65.445	-0.743	1.00	28.54
961	CB	ARG	287	2.009	66.390	-1.172	1.00	28.39
962	CG	ARG	287	0.672	66.025	-0.562	1.00	29.53
963	CD	ARG	287	-0.460	66.847	-1.130	1.00	29.36
964	NE	ARG	287	-1.743	66.209	-0.847	1.00	29.42
965	CZ	ARG	287	-2.922	66.686	-1.238	1.00	29.07
966	NH1	ARG	287	-2.990	67.821	-1.930	1.00	27.06
967	NH2	ARG	287	-4.033	66.014	-0.956	1.00	24.33
968	C	ARG	287	4.228	65.465	-1.773	1.00	28.07
969	O	ARG	287	4.327	64.584	-2.629	1.00	27.43
970	N	PHE	288	5.072	66.475	-1.686	1.00	28.65
971	CA	PHE	288	6.149	66.581	-2.639	1.00	30.74
972	CB	PHE	288	6.850	67.907	-2.474	1.00	33.46
973	CG	PHE	288	6.291	69.015	-3.335	1.00	37.89
974	CD1	PHE	288	5.658	70.109	-2.753	1.00	42.40
975	CD2	PHE	288	6.507	69.023	-4.706	1.00	40.19
976	CE1	PHE	288	5.262	71.203	-3.512	1.00	46.33
977	CE2	PHE	288	6.113	70.109	-5.487	1.00	45.03
978	CZ	PHE	288	5.490	71.207	-4.890	1.00	46.74
979	C	PHE	288	7.142	65.440	-2.532	1.00	29.68
980	O	PHE	288	7.812	65.114	-3.510	1.00	31.42
981	N	ASN	289	7.236	64.828	-1.355	1.00	28.29
982	CA	ASN	289	8.139	63.699	-1.164	1.00	26.35
983	CB	ASN	289	8.162	63.279	0.303	1.00	27.14

984	CG	ASN	289	9.208	62.217	0.584	1.00	27.91
985	OD1	ASN	289	8.928	61.016	0.544	1.00	28.10
986	ND2	ASN	289	10.433	62.658	0.847	1.00	26.93
987	C	ASN	289	7.700	62.513	-2.020	1.00	26.44
988	O	ASN	289	8.529	61.770	-2.537	1.00	27.45
989	N	THR	290	6.395	62.337	-2.181	1.00	25.31
990	CA	THR	290	5.907	61.227	-2.979	1.00	25.85
991	CB	THR	290	4.384	61.011	-2.800	1.00	27.70
992	OG1	THR	290	3.671	62.174	-3.247	1.00	27.96
993	CG2	THR	290	4.061	60.724	-1.335	1.00	24.95
994	C	THR	290	6.212	61.379	-4.461	1.00	26.07
995	O	THR	290	6.128	60.404	-5.215	1.00	24.53
996	N	VAL	291	6.534	62.592	-4.904	1.00	27.10
997	CA	VAL	291	6.873	62.754	-6.314	1.00	27.77
998	CB	VAL	291	6.076	63.885	-7.012	1.00	27.06
999	CG1	VAL	291	4.600	63.496	-7.097	1.00	26.44
1000	CG2	VAL	291	6.262	65.199	-6.283	1.00	27.18
1001	C	VAL	291	8.360	63.014	-6.445	1.00	29.18
1002	O	VAL	291	8.847	63.363	-7.522	1.00	29.91
1003	N	PHE	292	9.078	62.824	-5.337	1.00	30.55
1004	CA	PHE	292	10.528	63.006	-5.304	1.00	32.02
1005	CB	PHE	292	11.005	63.311	-3.891	1.00	31.56
1006	CG	PHE	292	12.491	63.472	-3.777	1.00	32.00
1007	CD1	PHE	292	13.145	64.500	-4.448	1.00	33.21
1008	CD2	PHE	292	13.237	62.632	-2.959	1.00	32.71
1009	CE1	PHE	292	14.519	64.699	-4.299	1.00	32.89
1010	CE2	PHE	292	14.616	62.822	-2.803	1.00	32.50
1011	CZ	PHE	292	15.255	63.860	-3.475	1.00	31.30
1012	C	PHE	292	11.212	61.738	-5.778	1.00	33.44
1013	O	PHE	292	10.752	60.630	-5.511	1.00	34.95
1014	N	ASN	293	12.319	61.909	-6.478	1.00	34.97
1015	CA	ASN	293	13.079	60.786	-6.998	1.00	37.05
1016	CB	ASN	293	13.114	60.879	-8.527	1.00	38.48
1017	CG	ASN	293	13.884	59.752	-9.169	1.00	41.16
1018	OD1	ASN	293	13.895	59.625	-10.390	1.00	43.01
1019	ND2	ASN	293	14.542	58.929	-8.356	1.00	42.39
1020	C	ASN	293	14.484	60.896	-6.398	1.00	38.22
1021	O	ASN	293	15.245	61.804	-6.743	1.00	36.98
1022	N	ALA	294	14.814	59.981	-5.490	1.00	39.59
1023	CA	ALA	294	16.114	59.987	-4.822	1.00	41.76
1024	CB	ALA	294	16.094	59.043	-3.627	1.00	40.19
1025	C	ALA	294	17.261	59.620	-5.761	1.00	43.28
1026	O	ALA	294	18.416	59.933	-5.486	1.00	43.11
1027	N	GLU	295	16.941	58.960	-6.867	1.00	44.98
1028	CA	GLU	295	17.956	58.581	-7.840	1.00	46.92
1029	CB	GLU	295	17.361	57.632	-8.881	1.00	49.77
1030	CG	GLU	295	16.828	56.343	-8.304	1.00	55.10
1031	CD	GLU	295	17.879	55.592	-7.516	1.00	58.35
1032	OE1	GLU	295	18.371	56.142	-6.504	1.00	61.01
1033	OE2	GLU	295	18.213	54.454	-7.913	1.00	60.18
1034	C	GLU	295	18.508	59.809	-8.554	1.00	46.47
1035	O	GLU	295	19.691	60.144	-8.431	1.00	45.54
1036	N	THR	296	17.629	60.473	-9.298	1.00	44.85
1037	CA	THR	296	17.989	61.654	-10.065	1.00	44.39
1038	CB	THR	296	17.070	61.797	-11.277	1.00	45.72

1039	OG1	THR	296	15.872	62.480	-10.889	1.00	48.92
1040	CG2	THR	296	16.686	60.427	-11.804	1.00	46.26
1041	C	THR	296	17.908	62.947	-9.257	1.00	43.36
1042	O	THR	296	18.158	64.030	-9.787	1.00	43.87
1043	N	GLY	297	17.554	62.840	-7.983	1.00	43.26
1044	CA	GLY	297	17.449	64.026	-7.145	1.00	42.33
1045	C	GLY	297	16.529	65.085	-7.730	1.00	41.54
1046	O	GLY	297	16.859	66.269	-7.727	1.00	40.23
1047	N	THR	298	15.370	64.663	-8.232	1.00	40.90
1048	CA	THR	298	14.417	65.597	-8.824	1.00	39.95
1049	CB	THR	298	14.480	65.549	-10.356	1.00	40.32
1050	OG1	THR	298	13.816	64.369	-10.828	1.00	40.79
1051	CG2	THR	298	15.925	65.524	-10.817	1.00	40.83
1052	C	THR	298	12.964	65.339	-8.410	1.00	39.33
1053	O	THR	298	12.587	64.217	-8.057	1.00	38.72
1054	N	TRP	299	12.153	66.393	-8.465	1.00	37.24
1055	CA	TRP	299	10.748	66.291	-8.120	1.00	35.84
1056	CB	TRP	299	10.360	67.433	-7.192	1.00	36.09
1057	CG	TRP	299	10.922	67.293	-5.818	1.00	37.82
1058	CD2	TRP	299	12.122	67.893	-5.322	1.00	38.33
1059	CE2	TRP	299	12.276	67.478	-3.975	1.00	38.53
1060	CE3	TRP	299	13.086	68.739	-5.884	1.00	37.90
1061	CD1	TRP	299	10.405	66.550	-4.781	1.00	37.23
1062	NE1	TRP	299	11.214	66.664	-3.668	1.00	36.02
1063	CZ2	TRP	299	13.360	67.885	-3.184	1.00	38.17
1064	CZ3	TRP	299	14.163	69.144	-5.096	1.00	38.48
1065	CH2	TRP	299	14.290	68.716	-3.761	1.00	38.37
1066	C	TRP	299	9.873	66.321	-9.362	1.00	35.79
1067	O	TRP	299	9.675	67.378	-9.964	1.00	35.85
1068	N	GLU	300	9.361	65.156	-9.749	1.00	34.45
1069	CA	GLU	300	8.486	65.041	-10.911	1.00	34.50
1070	CB	GLU	300	8.237	63.573	-11.294	1.00	38.87
1071	CG	GLU	300	9.153	62.956	-12.340	1.00	44.29
1072	CD	GLU	300	10.527	62.603	-11.796	1.00	47.59
1073	OE1	GLU	300	11.163	61.677	-12.353	1.00	47.72
1074	OE2	GLU	300	10.973	63.255	-10.822	1.00	50.29
1075	C	GLU	300	7.129	65.640	-10.594	1.00	31.52
1076	O	GLU	300	6.373	65.061	-9.822	1.00	31.47
1077	N	CYS	301	6.808	66.777	-11.195	1.00	28.82
1078	CA	CYS	301	5.508	67.387	-10.967	1.00	28.63
1079	CB	CYS	301	5.665	68.747	-10.293	1.00	26.70
1080	SG	CYS	301	6.609	68.698	-8.765	1.00	27.52
1081	C	CYS	301	4.781	67.551	-12.294	1.00	29.62
1082	O	CYS	301	4.592	68.671	-12.772	1.00	30.22
1083	N	GLY	302	4.367	66.437	-12.890	1.00	29.81
1084	CA	GLY	302	3.673	66.512	-14.163	1.00	31.07
1085	C	GLY	302	4.648	66.864	-15.267	1.00	32.36
1086	O	GLY	302	5.721	66.266	-15.350	1.00	33.42
1087	N	ARG	303	4.293	67.833	-16.109	1.00	33.44
1088	CA	ARG	303	5.177	68.245	-17.201	1.00	35.64
1089	CB	ARG	303	4.427	69.112	-18.224	1.00	37.56
1090	CG	ARG	303	3.036	68.639	-18.619	1.00	41.48
1091	CD	ARG	303	2.986	67.925	-19.967	1.00	43.33
1092	NE	ARG	303	1.597	67.796	-20.415	1.00	46.76
1093	CZ	ARG	303	1.194	67.073	-21.459	1.00	49.83

1094	NH1	ARG	303	2.074	66.392	-22.187	1.00	50.70
1095	NH2	ARG	303	-0.097	67.032	-21.779	1.00	50.75
1096	C	ARG	303	6.324	69.071	-16.619	1.00	35.57
1097	O	ARG	303	7.301	69.382	-17.299	1.00	36.63
1098	N	LEU	304	6.200	69.425	-15.350	1.00	35.86
1099	CA	LEU	304	7.210	70.228	-14.689	1.00	36.28
1100	CB	LEU	304	6.511	71.224	-13.774	1.00	35.92
1101	CG	LEU	304	7.004	72.664	-13.808	1.00	39.08
1102	CD1	LEU	304	5.939	73.579	-13.184	1.00	39.61
1103	CD2	LEU	304	8.338	72.762	-13.066	1.00	41.12
1104	C	LEU	304	8.205	69.371	-13.896	1.00	36.47
1105	O	LEU	304	7.876	68.277	-13.436	1.00	36.05
1106	N	SER	305	9.428	69.869	-13.759	1.00	37.29
1107	CA	SER	305	10.470	69.167	-13.015	1.00	37.59
1108	CB	SER	305	11.258	68.236	-13.926	1.00	37.97
1109	OG	SER	305	10.583	67.005	-14.070	1.00	42.64
1110	C	SER	305	11.437	70.104	-12.321	1.00	37.22
1111	O	SER	305	11.810	71.149	-12.854	1.00	36.04
1112	N	TYR	306	11.823	69.728	-11.112	1.00	38.23
1113	CA	TYR	306	12.765	70.515	-10.341	1.00	40.56
1114	CB	TYR	306	12.130	71.016	-9.050	1.00	40.37
1115	CG	TYR	306	10.881	71.822	-9.254	1.00	40.12
1116	CD1	TYR	306	9.641	71.205	-9.393	1.00	38.75
1117	CE1	TYR	306	8.483	71.954	-9.575	1.00	39.38
1118	CD2	TYR	306	10.934	73.212	-9.305	1.00	41.30
1119	CE2	TYR	306	9.782	73.970	-9.487	1.00	40.89
1120	CZ	TYR	306	8.564	73.336	-9.618	1.00	40.00
1121	OH	TYR	306	7.435	74.097	-9.787	1.00	41.04
1122	C	TYR	306	13.921	69.598	-10.012	1.00	41.94
1123	O	TYR	306	13.814	68.749	-9.126	1.00	41.85
1124	N	CYS	307	15.020	69.751	-10.740	1.00	44.54
1125	CA	CYS	307	16.186	68.917	-10.504	1.00	47.11
1126	CB	CYS	307	16.889	68.590	-11.822	1.00	48.12
1127	SG	CYS	307	18.342	67.524	-11.637	1.00	53.21
1128	C	CYS	307	17.142	69.624	-9.562	1.00	46.97
1129	O	CYS	307	17.394	70.814	-9.701	1.00	45.58
1130	N	LEU	308	17.647	68.878	-8.587	1.00	50.01
1131	CA	LEU	308	18.585	69.405	-7.603	1.00	54.07
1132	CB	LEU	308	18.693	68.431	-6.427	1.00	51.75
1133	CG	LEU	308	18.614	68.978	-5.001	1.00	51.56
1134	CD1	LEU	308	17.433	69.914	-4.838	1.00	50.42
1135	CD2	LEU	308	18.486	67.811	-4.043	1.00	51.74
1136	C	LEU	308	19.938	69.560	-8.288	1.00	58.05
1137	O	LEU	308	20.423	68.628	-8.936	1.00	57.64
1138	N	GLU	309	20.533	70.741	-8.170	1.00	63.36
1139	CA	GLU	309	21.827	71.004	-8.787	1.00	69.50
1140	CB	GLU	309	22.159	72.500	-8.703	1.00	70.80
1141	CG	GLU	309	21.420	73.371	-9.716	1.00	72.19
1142	CD	GLU	309	21.856	73.102	-11.154	1.00	74.35
1143	OE1	GLU	309	22.720	72.220	-11.367	1.00	74.21
1144	OE2	GLU	309	21.333	73.771	-12.074	1.00	75.21
1145	C	GLU	309	22.929	70.180	-8.121	1.00	72.84
1146	O	GLU	309	23.958	70.715	-7.718	1.00	73.08
1147	N	ASP	310	22.695	68.872	-8.008	1.00	77.53
1148	CA	ASP	310	23.647	67.940	-7.400	1.00	80.81

1149	CB	ASP	310	22.921	66.707	-6.852	1.00	80.35
1150	C	ASP	310	24.664	67.505	-8.444	1.00	83.36
1151	O	ASP	310	24.391	67.561	-9.646	1.00	84.05
1152	N	THR	311	25.829	67.056	-7.983	1.00	86.01
1153	CA	THR	311	26.897	66.642	-8.889	1.00	88.11
1154	CB	THR	311	26.502	65.389	-9.701	1.00	87.28
1155	C	THR	311	27.146	67.809	-9.844	1.00	89.36
1156	O	THR	311	27.672	67.634	-10.943	1.00	89.37
1157	N	ALA	312	26.746	69.000	-9.396	1.00	90.86
1158	CA	ALA	312	26.893	70.240	-10.154	1.00	92.19
1159	CB	ALA	312	25.626	70.507	-10.974	1.00	91.96
1160	C	ALA	312	27.148	71.398	-9.183	1.00	92.90
1161	O	ALA	312	28.154	72.106	-9.294	1.00	93.06
1162	N	GLY	313	26.232	71.577	-8.233	1.00	92.79
1163	CA	GLY	313	26.362	72.636	-7.247	1.00	92.69
1164	C	GLY	313	27.157	72.182	-6.035	1.00	92.78
1165	O	GLY	313	28.386	72.284	-6.014	1.00	93.48
1166	N	GLY	314	26.458	71.686	-5.018	1.00	92.11
1167	CA	GLY	314	27.136	71.215	-3.823	1.00	91.35
1168	C	GLY	314	27.735	69.840	-4.060	1.00	90.85
1169	O	GLY	314	28.429	69.627	-5.057	1.00	91.34
1170	N	PHE	315	27.478	68.910	-3.143	1.00	89.85
1171	CA	PHE	315	27.982	67.539	-3.255	1.00	88.10
1172	CB	PHE	315	29.375	67.433	-2.607	1.00	87.89
1173	CG	PHE	315	30.380	68.414	-3.173	1.00	87.72
1174	CD1	PHE	315	30.452	69.719	-2.687	1.00	87.77
1175	CD2	PHE	315	31.194	68.060	-4.248	1.00	87.87
1176	CE1	PHE	315	31.315	70.658	-3.267	1.00	87.63
1177	CE2	PHE	315	32.059	68.993	-4.837	1.00	87.91
1178	CZ	PHE	315	32.117	70.294	-4.345	1.00	87.66
1179	C	PHE	315	26.973	66.584	-2.599	1.00	86.68
1180	O	PHE	315	26.440	66.877	-1.531	1.00	86.63
1181	N	GLN	316	26.706	65.455	-3.251	1.00	85.03
1182	CA	GLN	316	25.732	64.464	-2.773	1.00	83.31
1183	CB	GLN	316	25.729	63.252	-3.709	1.00	83.45
1184	C	GLN	316	25.845	63.972	-1.326	1.00	82.19
1185	O	GLN	316	25.129	63.050	-0.925	1.00	80.95
1186	N	GLN	317	26.746	64.572	-0.553	1.00	81.05
1187	CA	GLN	317	26.927	64.219	0.855	1.00	78.97
1188	CB	GLN	317	28.366	63.763	1.122	1.00	79.63
1189	C	GLN	317	26.627	65.484	1.650	1.00	76.87
1190	O	GLN	317	26.388	65.443	2.859	1.00	77.12
1191	N	LEU	318	26.637	66.603	0.929	1.00	74.30
1192	CA	LEU	318	26.368	67.932	1.468	1.00	71.67
1193	CB	LEU	318	26.890	68.986	0.475	1.00	71.47
1194	CG	LEU	318	27.050	70.477	0.801	1.00	71.19
1195	CD1	LEU	318	28.154	71.041	-0.086	1.00	70.22
1196	CD2	LEU	318	25.751	71.241	0.586	1.00	69.44
1197	C	LEU	318	24.860	68.099	1.707	1.00	69.87
1198	O	LEU	318	24.278	69.140	1.403	1.00	69.31
1199	N	LEU	319	24.236	67.045	2.234	1.00	67.66
1200	CA	LEU	319	22.809	67.048	2.550	1.00	64.15
1201	CB	LEU	319	22.250	65.627	2.601	1.00	63.02
1202	CG	LEU	319	22.141	64.812	1.316	1.00	63.39
1203	CD1	LEU	319	23.458	64.819	0.566	1.00	62.52

1204	CD2	LEU	319	21.726	63.392	1.673	1.00	62.35
1205	C	LEU	319	22.686	67.662	3.929	1.00	62.88
1206	O	LEU	319	21.840	67.261	4.728	1.00	62.86
1207	N	LEU	320	23.563	68.615	4.216	1.00	60.82
1208	CA	LEU	320	23.534	69.287	5.500	1.00	58.59
1209	CB	LEU	320	24.757	70.202	5.662	1.00	59.36
1210	CG	LEU	320	26.113	69.519	5.894	1.00	59.11
1211	CD1	LEU	320	25.959	68.436	6.957	1.00	58.60
1212	CD2	LEU	320	26.622	68.908	4.605	1.00	59.72
1213	C	LEU	320	22.246	70.090	5.601	1.00	56.13
1214	O	LEU	320	21.823	70.460	6.692	1.00	55.60
1215	N	GLU	321	21.628	70.350	4.452	1.00	54.36
1216	CA	GLU	321	20.373	71.089	4.405	1.00	51.96
1217	CB	GLU	321	20.020	71.447	2.960	1.00	52.01
1218	CG	GLU	321	18.894	72.451	2.847	1.00	53.96
1219	CD	GLU	321	19.252	73.791	3.455	1.00	55.08
1220	OE1	GLU	321	18.329	74.596	3.703	1.00	56.88
1221	OE2	GLU	321	20.454	74.047	3.675	1.00	54.69
1222	C	GLU	321	19.307	70.179	5.012	1.00	49.92
1223	O	GLU	321	18.996	69.110	4.479	1.00	48.40
1224	N	PRO	322	18.744	70.591	6.152	1.00	48.78
1225	CD	PRO	322	18.957	71.878	6.833	1.00	47.26
1226	CA	PRO	322	17.720	69.811	6.845	1.00	47.76
1227	CB	PRO	322	17.174	70.801	7.869	1.00	47.73
1228	CG	PRO	322	18.376	71.600	8.216	1.00	47.28
1229	C	PRO	322	16.629	69.261	5.942	1.00	47.10
1230	O	PRO	322	16.317	68.069	5.989	1.00	48.08
1231	N	MET	323	16.056	70.116	5.104	1.00	45.34
1232	CA	MET	323	14.976	69.655	4.259	1.00	44.05
1233	CB	MET	323	14.307	70.816	3.563	1.00	44.58
1234	CG	MET	323	13.097	70.333	2.844	1.00	46.85
1235	SD	MET	323	11.694	70.427	3.935	1.00	47.74
1236	CE	MET	323	11.702	68.792	4.625	1.00	47.08
1237	C	MET	323	15.329	68.598	3.220	1.00	42.42
1238	O	MET	323	14.547	67.679	2.994	1.00	41.60
1239	N	LEU	324	16.483	68.723	2.575	1.00	41.40
1240	CA	LEU	324	16.867	67.731	1.587	1.00	41.52
1241	CB	LEU	324	18.161	68.132	0.875	1.00	44.57
1242	CG	LEU	324	18.089	69.222	-0.207	1.00	47.29
1243	CD1	LEU	324	16.943	68.912	-1.175	1.00	48.25
1244	CD2	LEU	324	17.884	70.581	0.430	1.00	46.68
1245	C	LEU	324	17.058	66.409	2.305	1.00	41.54
1246	O	LEU	324	16.565	65.366	1.864	1.00	41.10
1247	N	LYS	325	17.769	66.461	3.426	1.00	40.55
1248	CA	LYS	325	18.015	65.270	4.223	1.00	40.14
1249	CB	LYS	325	18.735	65.641	5.523	1.00	42.78
1250	CG	LYS	325	19.276	64.450	6.294	1.00	46.67
1251	CD	LYS	325	19.970	64.880	7.589	1.00	51.48
1252	CE	LYS	325	20.585	63.683	8.325	1.00	52.26
1253	NZ	LYS	325	21.122	64.074	9.657	1.00	53.91
1254	C	LYS	325	16.674	64.621	4.544	1.00	37.65
1255	O	LYS	325	16.539	63.391	4.527	1.00	36.76
1256	N	PHE	326	15.682	65.455	4.837	1.00	33.98
1257	CA	PHE	326	14.349	64.957	5.152	1.00	32.32
1258	CB	PHE	326	13.419	66.117	5.503	1.00	30.44

1259	CG	PHE	326	12.044	65.687	5.900	1.00	28.49
1260	CD1	PHE	326	11.791	65.229	7.183	1.00	28.47
1261	CD2	PHE	326	10.993	65.737	4.984	1.00	29.48
1262	CE1	PHE	326	10.505	64.826	7.554	1.00	30.80
1263	CE2	PHE	326	9.708	65.339	5.338	1.00	27.50
1264	CZ	PHE	326	9.460	64.884	6.625	1.00	29.82
1265	C	PHE	326	13.759	64.175	3.976	1.00	30.42
1266	O	PHE	326	13.279	63.057	4.139	1.00	31.78
1267	N	HIS	327	13.799	64.758	2.786	1.00	27.63
1268	CA	HIS	327	13.254	64.080	1.630	1.00	27.68
1269	CB	HIS	327	13.253	65.026	0.422	1.00	26.07
1270	CG	HIS	327	12.076	65.954	0.382	1.00	25.45
1271	CD2	HIS	327	11.972	67.280	0.645	1.00	26.41
1272	ND1	HIS	327	10.799	65.529	0.075	1.00	25.84
1273	CE1	HIS	327	9.961	66.549	0.152	1.00	22.17
1274	NE2	HIS	327	10.647	67.623	0.497	1.00	24.58
1275	C	HIS	327	13.965	62.754	1.305	1.00	28.24
1276	O	HIS	327	13.303	61.744	1.068	1.00	27.91
1277	N	TYR	328	15.295	62.731	1.291	1.00	29.44
1278	CA	TYR	328	15.986	61.473	0.998	1.00	31.45
1279	CB	TYR	328	17.490	61.671	0.911	1.00	30.86
1280	CG	TYR	328	17.940	62.287	-0.386	1.00	33.71
1281	CD1	TYR	328	18.241	63.645	-0.472	1.00	34.16
1282	CE1	TYR	328	18.650	64.209	-1.672	1.00	34.37
1283	CD2	TYR	328	18.057	61.512	-1.535	1.00	33.47
1284	CE2	TYR	328	18.459	62.067	-2.736	1.00	34.02
1285	CZ	TYR	328	18.753	63.413	-2.797	1.00	34.35
1286	OH	TYR	328	19.142	63.964	-3.990	1.00	37.80
1287	C	TYR	328	15.697	60.445	2.076	1.00	33.10
1288	O	TYR	328	15.438	59.272	1.793	1.00	33.50
1289	N	MET	329	15.741	60.895	3.323	1.00	34.80
1290	CA	MET	329	15.482	60.011	4.444	1.00	36.18
1291	CB	MET	329	15.766	60.740	5.755	1.00	37.18
1292	CG	MET	329	17.220	60.619	6.187	1.00	40.38
1293	SD	MET	329	17.538	61.382	7.789	1.00	46.12
1294	CE	MET	329	16.147	60.660	8.773	1.00	43.53
1295	C	MET	329	14.068	59.434	4.444	1.00	35.91
1296	O	MET	329	13.895	58.236	4.692	1.00	36.25
1297	N	LEU	330	13.066	60.273	4.164	1.00	34.27
1298	CA	LEU	330	11.687	59.809	4.135	1.00	33.17
1299	CB	LEU	330	10.698	60.984	4.109	1.00	30.57
1300	CG	LEU	330	9.202	60.609	4.128	1.00	28.18
1301	CD1	LEU	330	8.918	59.589	5.214	1.00	27.72
1302	CD2	LEU	330	8.362	61.840	4.348	1.00	27.96
1303	C	LEU	330	11.459	58.918	2.922	1.00	34.37
1304	O	LEU	330	10.763	57.904	3.010	1.00	34.32
1305	N	LYS	331	12.050	59.291	1.793	1.00	34.21
1306	CA	LYS	331	11.896	58.498	0.586	1.00	35.75
1307	CB	LYS	331	12.610	59.176	-0.587	1.00	36.14
1308	CG	LYS	331	11.723	59.429	-1.816	1.00	36.00
1309	CD	LYS	331	11.202	58.133	-2.395	1.00	36.30
1310	CE	LYS	331	10.371	58.357	-3.653	1.00	36.58
1311	NZ	LYS	331	9.140	59.136	-3.396	1.00	34.88
1312	C	LYS	331	12.493	57.120	0.825	1.00	37.43
1313	O	LYS	331	12.001	56.116	0.310	1.00	37.00

1314	N	LYS	332	13.549	57.075	1.629	1.00	39.12
1315	CA	LYS	332	14.226	55.819	1.914	1.00	40.90
1316	CB	LYS	332	15.490	56.075	2.741	1.00	43.72
1317	CG	LYS	332	16.771	55.546	2.094	1.00	46.03
1318	CD	LYS	332	16.750	54.024	1.984	1.00	49.52
1319	CE	LYS	332	16.850	53.360	3.359	1.00	50.49
1320	NZ	LYS	332	18.186	53.576	3.993	1.00	51.08
1321	C	LYS	332	13.341	54.799	2.620	1.00	40.65
1322	O	LYS	332	13.465	53.602	2.383	1.00	42.59
1323	N	LEU	333	12.445	55.258	3.481	1.00	38.72
1324	CA	LEU	333	11.571	54.333	4.191	1.00	37.51
1325	CB	LEU	333	10.762	55.083	5.251	1.00	37.44
1326	CG	LEU	333	11.581	55.907	6.252	1.00	38.27
1327	CD1	LEU	333	10.674	56.906	6.958	1.00	38.07
1328	CD2	LEU	333	12.267	54.990	7.250	1.00	37.10
1329	C	LEU	333	10.617	53.601	3.245	1.00	36.94
1330	O	LEU	333	10.034	52.578	3.616	1.00	37.20
1331	N	GLN	334	10.466	54.114	2.025	1.00	36.34
1332	CA	GLN	334	9.556	53.513	1.038	1.00	36.67
1333	CB	GLN	334	10.072	52.142	0.582	1.00	39.18
1334	CG	GLN	334	11.362	52.183	-0.218	1.00	44.20
1335	CD	GLN	334	11.729	50.827	-0.791	1.00	46.80
1336	OE1	GLN	334	11.064	50.312	-1.697	1.00	46.66
1337	NE2	GLN	334	12.792	50.234	-0.259	1.00	48.21
1338	C	GLN	334	8.137	53.341	1.586	1.00	33.92
1339	O	GLN	334	7.545	52.270	1.471	1.00	31.98
1340	N	LEU	335	7.587	54.397	2.169	1.00	32.76
1341	CA	LEU	335	6.253	54.314	2.752	1.00	32.40
1342	CB	LEU	335	5.930	55.594	3.530	1.00	29.76
1343	CG	LEU	335	6.856	56.027	4.669	1.00	30.25
1344	CD1	LEU	335	6.257	57.255	5.368	1.00	28.30
1345	CD2	LEU	335	7.042	54.886	5.662	1.00	28.50
1346	C	LEU	335	5.164	54.085	1.716	1.00	32.26
1347	O	LEU	335	5.345	54.373	0.534	1.00	29.50
1348	N	HIS	336	4.033	53.564	2.183	1.00	34.06
1349	CA	HIS	336	2.868	53.320	1.331	1.00	34.84
1350	CB	HIS	336	2.003	52.185	1.892	1.00	36.51
1351	CG	HIS	336	2.683	50.852	1.922	1.00	37.66
1352	CD2	HIS	336	3.922	50.470	1.533	1.00	37.54
1353	ND1	HIS	336	2.070	49.720	2.416	1.00	37.21
1354	CE1	HIS	336	2.904	48.699	2.330	1.00	38.62
1355	NE2	HIS	336	4.034	49.127	1.798	1.00	38.49
1356	C	HIS	336	2.022	54.596	1.315	1.00	34.82
1357	O	HIS	336	2.211	55.496	2.155	1.00	33.88
1358	N	GLU	337	1.087	54.662	0.369	1.00	32.20
1359	CA	GLU	337	0.201	55.809	0.253	1.00	31.43
1360	CB	GLU	337	-0.779	55.610	-0.904	1.00	32.58
1361	CG	GLU	337	-0.124	55.491	-2.281	1.00	34.45
1362	CD	GLU	337	0.186	56.837	-2.921	1.00	35.05
1363	OE1	GLU	337	0.125	57.868	-2.211	1.00	34.61
1364	OE2	GLU	337	0.498	56.859	-4.137	1.00	34.56
1365	C	GLU	337	-0.575	56.000	1.551	1.00	31.12
1366	O	GLU	337	-0.730	57.124	2.021	1.00	31.97
1367	N	GLU	338	-1.046	54.904	2.140	1.00	30.96
1368	CA	GLU	338	-1.820	54.987	3.382	1.00	31.94

1369	CB	GLU	338	-2.310	53.603	3.831	1.00	32.40
1370	CG	GLU	338	-3.191	52.899	2.829	1.00	33.98
1371	CD	GLU	338	-2.418	52.382	1.632	1.00	37.65
1372	OE1	GLU	338	-3.062	51.827	0.716	1.00	40.59
1373	OE2	GLU	338	-1.171	52.518	1.603	1.00	38.76
1374	C	GLU	338	-1.004	55.601	4.497	1.00	31.26
1375	O	GLU	338	-1.519	56.382	5.296	1.00	31.58
1376	N	GLU	339	0.269	55.230	4.552	1.00	31.23
1377	CA	GLU	339	1.168	55.745	5.565	1.00	31.21
1378	CB	GLU	339	2.496	54.973	5.509	1.00	31.87
1379	CG	GLU	339	2.410	53.584	6.167	1.00	30.48
1380	CD	GLU	339	3.533	52.631	5.768	1.00	30.53
1381	OE1	GLU	339	3.620	51.540	6.377	1.00	27.16
1382	OE2	GLU	339	4.320	52.958	4.845	1.00	32.97
1383	C	GLU	339	1.360	57.250	5.347	1.00	30.76
1384	O	GLU	339	1.325	58.046	6.289	1.00	29.59
1385	N	TYR	340	1.530	57.648	4.099	1.00	29.64
1386	CA	TYR	340	1.688	59.062	3.820	1.00	30.11
1387	CB	TYR	340	1.970	59.287	2.338	1.00	28.18
1388	CG	TYR	340	3.423	59.198	1.980	1.00	28.53
1389	CD1	TYR	340	3.938	58.070	1.335	1.00	29.29
1390	CE1	TYR	340	5.273	58.013	0.946	1.00	27.77
1391	CD2	TYR	340	4.286	60.265	2.235	1.00	28.14
1392	CE2	TYR	340	5.621	60.219	1.850	1.00	27.25
1393	CZ	TYR	340	6.105	59.093	1.204	1.00	27.82
1394	OH	TYR	340	7.415	59.062	0.791	1.00	29.96
1395	C	TYR	340	0.446	59.859	4.237	1.00	30.96
1396	O	TYR	340	0.573	60.936	4.829	1.00	32.47
1397	N	VAL	341	-0.745	59.347	3.922	1.00	29.56
1398	CA	VAL	341	-1.989	60.037	4.281	1.00	28.92
1399	CB	VAL	341	-3.241	59.309	3.761	1.00	27.33
1400	CG1	VAL	341	-4.453	60.154	4.003	1.00	25.94
1401	CG2	VAL	341	-3.104	59.025	2.300	1.00	30.42
1402	C	VAL	341	-2.130	60.129	5.798	1.00	29.62
1403	O	VAL	341	-2.566	61.154	6.338	1.00	28.89
1404	N	LEU	342	-1.769	59.047	6.479	1.00	28.26
1405	CA	LEU	342	-1.853	59.018	7.928	1.00	28.25
1406	CB	LEU	342	-1.559	57.604	8.442	1.00	27.91
1407	CG	LEU	342	-2.785	56.694	8.563	1.00	25.84
1408	CD1	LEU	342	-2.348	55.271	8.833	1.00	24.76
1409	CD2	LEU	342	-3.685	57.199	9.687	1.00	24.87
1410	C	LEU	342	-0.909	60.037	8.575	1.00	27.93
1411	O	LEU	342	-1.312	60.770	9.479	1.00	24.90
1412	N	MET	343	0.340	60.099	8.114	1.00	28.60
1413	CA	MET	343	1.255	61.053	8.709	1.00	29.22
1414	CB	MET	343	2.731	60.664	8.482	1.00	28.71
1415	CG	MET	343	3.255	60.640	7.074	1.00	29.84
1416	SD	MET	343	4.936	59.888	7.019	1.00	33.27
1417	CE	MET	343	6.012	61.293	7.303	1.00	28.58
1418	C	MET	343	0.940	62.453	8.226	1.00	30.35
1419	O	MET	343	1.562	63.408	8.662	1.00	34.05
1420	N	GLN	344	-0.041	62.574	7.332	1.00	30.98
1421	CA	GLN	344	-0.501	63.884	6.862	1.00	30.24
1422	CB	GLN	344	-1.075	63.810	5.448	1.00	28.33
1423	CG	GLN	344	-0.088	63.959	4.307	1.00	28.64

1424	CD	GLN	344	-0.786	63.928	2.958	1.00	28.44
1425	OE1	GLN	344	-0.201	63.538	1.937	1.00	26.65
1426	NE2	GLN	344	-2.055	64.340	2.948	1.00	25.63
1427	C	GLN	344	-1.635	64.278	7.824	1.00	31.48
1428	O	GLN	344	-1.915	65.459	8.032	1.00	31.60
1429	N	ALA	345	-2.292	63.268	8.397	1.00	31.49
1430	CA	ALA	345	-3.380	63.495	9.345	1.00	32.16
1431	CB	ALA	345	-4.215	62.226	9.515	1.00	32.71
1432	C	ALA	345	-2.790	63.905	10.682	1.00	31.48
1433	O	ALA	345	-3.209	64.895	11.290	1.00	32.08
1434	N	ILE	346	-1.810	63.132	11.134	1.00	29.84
1435	CA	ILE	346	-1.154	63.410	12.394	1.00	27.87
1436	CB	ILE	346	-0.030	62.422	12.634	1.00	26.76
1437	CG2	ILE	346	0.714	62.771	13.918	1.00	24.05
1438	CG1	ILE	346	-0.621	61.017	12.698	1.00	25.01
1439	CD1	ILE	346	0.407	59.923	12.747	1.00	28.26
1440	C	ILE	346	-0.601	64.825	12.385	1.00	29.29
1441	O	ILE	346	-0.731	65.560	13.372	1.00	30.78
1442	N	SER	347	0.007	65.221	11.272	1.00	27.42
1443	CA	SER	347	0.539	66.569	11.192	1.00	26.56
1444	CB	SER	347	1.349	66.757	9.914	1.00	25.97
1445	OG	SER	347	1.716	68.118	9.767	1.00	25.49
1446	C	SER	347	-0.588	67.598	11.228	1.00	26.93
1447	O	SER	347	-0.531	68.559	11.997	1.00	27.11
1448	N	LEU	348	-1.613	67.392	10.403	1.00	25.68
1449	CA	LEU	348	-2.739	68.325	10.334	1.00	27.35
1450	CB	LEU	348	-3.782	67.842	9.319	1.00	27.62
1451	CG	LEU	348	-5.007	68.749	9.149	1.00	27.62
1452	CD1	LEU	348	-4.632	69.981	8.351	1.00	27.49
1453	CD2	LEU	348	-6.115	67.995	8.437	1.00	28.51
1454	C	LEU	348	-3.417	68.535	11.681	1.00	28.45
1455	O	LEU	348	-3.749	69.664	12.045	1.00	25.93
1456	N	PHE	349	-3.640	67.438	12.403	1.00	31.00
1457	CA	PHE	349	-4.272	67.508	13.716	1.00	34.64
1458	CB	PHE	349	-5.238	66.328	13.933	1.00	34.22
1459	CG	PHE	349	-6.392	66.289	12.962	1.00	34.97
1460	CD1	PHE	349	-6.385	65.411	11.880	1.00	35.03
1461	CD2	PHE	349	-7.484	67.136	13.120	1.00	35.09
1462	CE1	PHE	349	-7.450	65.376	10.968	1.00	34.42
1463	CE2	PHE	349	-8.552	67.108	12.211	1.00	34.82
1464	CZ	PHE	349	-8.531	66.225	11.135	1.00	33.70
1465	C	PHE	349	-3.203	67.518	14.818	1.00	37.57
1466	O	PHE	349	-3.024	66.541	15.557	1.00	37.11
1467	N	SER	350	-2.474	68.627	14.896	1.00	39.69
1468	CA	SER	350	-1.443	68.799	15.905	1.00	42.41
1469	CB	SER	350	-0.198	69.453	15.309	1.00	41.46
1470	OG	SER	350	0.458	68.559	14.430	1.00	43.31
1471	C	SER	350	-2.063	69.703	16.947	1.00	44.41
1472	O	SER	350	-2.345	70.872	16.680	1.00	45.59
1473	N	PRO	351	-2.294	69.170	18.152	1.00	45.82
1474	CD	PRO	351	-1.910	67.824	18.614	1.00	45.34
1475	CA	PRO	351	-2.899	69.944	19.235	1.00	46.70
1476	CB	PRO	351	-3.109	68.896	20.322	1.00	45.62
1477	CG	PRO	351	-1.951	67.974	20.113	1.00	45.26
1478	C	PRO	351	-2.051	71.114	19.707	1.00	48.66

1479	O	PRO	351	-2.580	72.119	20.192	1.00	49.28
1480	N	ASP	352	-0.737	70.994	19.545	1.00	50.51
1481	CA	ASP	352	0.165	72.042	20.001	1.00	51.66
1482	CB	ASP	352	1.434	71.418	20.591	1.00	53.36
1483	CG	ASP	352	2.145	70.505	19.613	1.00	55.03
1484	OD1	ASP	352	3.055	69.760	20.046	1.00	54.46
1485	OD2	ASP	352	1.794	70.540	18.415	1.00	55.80
1486	C	ASP	352	0.538	73.104	18.981	1.00	51.49
1487	O	ASP	352	1.699	73.481	18.875	1.00	52.72
1488	N	ARG	353	-0.446	73.581	18.228	1.00	51.00
1489	CA	ARG	353	-0.201	74.639	17.258	1.00	51.20
1490	CB	ARG	353	-0.697	74.258	15.858	1.00	49.58
1491	CG	ARG	353	0.155	73.214	15.150	1.00	48.32
1492	CD	ARG	353	-0.207	73.104	13.676	1.00	47.64
1493	NE	ARG	353	0.394	71.941	13.017	1.00	45.87
1494	CZ	ARG	353	1.702	71.737	12.881	1.00	44.56
1495	NH1	ARG	353	2.568	72.622	13.358	1.00	43.10
1496	NH2	ARG	353	2.146	70.644	12.272	1.00	42.08
1497	C	ARG	353	-0.967	75.844	17.762	1.00	52.31
1498	O	ARG	353	-2.154	75.752	18.045	1.00	53.67
1499	N	PRO	354	-0.296	76.996	17.877	1.00	53.28
1500	CD	PRO	354	1.042	77.245	17.322	1.00	52.61
1501	CA	PRO	354	-0.894	78.251	18.353	1.00	54.12
1502	CB	PRO	354	0.114	79.312	17.895	1.00	53.64
1503	CG	PRO	354	0.879	78.633	16.786	1.00	52.56
1504	C	PRO	354	-2.323	78.579	17.904	1.00	54.55
1505	O	PRO	354	-2.687	78.381	16.744	1.00	54.94
1506	N	GLY	355	-3.124	79.080	18.846	1.00	55.02
1507	CA	GLY	355	-4.498	79.460	18.558	1.00	55.49
1508	C	GLY	355	-5.536	78.351	18.550	1.00	56.27
1509	O	GLY	355	-6.721	78.615	18.335	1.00	57.62
1510	N	VAL	356	-5.106	77.116	18.781	1.00	55.29
1511	CA	VAL	356	-6.020	75.981	18.780	1.00	55.29
1512	CB	VAL	356	-5.237	74.638	18.779	1.00	55.13
1513	CG1	VAL	356	-6.190	73.469	18.983	1.00	54.19
1514	CG2	VAL	356	-4.486	74.474	17.466	1.00	53.37
1515	C	VAL	356	-6.962	76.018	19.979	1.00	56.09
1516	O	VAL	356	-6.520	76.038	21.128	1.00	56.92
1517	N	GLN	357	-8.263	76.026	19.697	1.00	56.52
1518	CA	GLN	357	-9.292	76.059	20.735	1.00	56.33
1519	CB	GLN	357	-10.613	76.562	20.156	1.00	58.54
1520	CG	GLN	357	-10.537	77.967	19.586	1.00	62.76
1521	CD	GLN	357	-9.955	78.960	20.572	1.00	64.33
1522	OE1	GLN	357	-10.442	79.086	21.695	1.00	65.77
1523	NE2	GLN	357	-8.907	79.672	20.157	1.00	65.16
1524	C	GLN	357	-9.513	74.690	21.367	1.00	55.24
1525	O	GLN	357	-9.000	74.418	22.451	1.00	54.57
1526	N	LEU	358	-10.282	73.829	20.699	1.00	53.55
1527	CA	LEU	358	-10.532	72.498	21.241	1.00	52.75
1528	CB	LEU	358	-11.734	71.833	20.544	1.00	51.55
1529	CG	LEU	358	-13.142	72.335	20.927	1.00	50.47
1530	CD1	LEU	358	-14.202	71.397	20.357	1.00	49.39
1531	CD2	LEU	358	-13.289	72.402	22.444	1.00	49.10
1532	C	LEU	358	-9.290	71.622	21.122	1.00	51.67
1533	O	LEU	358	-9.294	70.607	20.435	1.00	50.15

1534	N	HIS	359	-8.230	72.030	21.814	1.00	52.68
1535	CA	HIS	359	-6.961	71.310	21.801	1.00	55.25
1536	CB	HIS	359	-5.884	72.120	22.530	1.00	57.48
1537	CG	HIS	359	-6.278	72.542	23.909	1.00	61.13
1538	CD2	HIS	359	-6.670	73.745	24.396	1.00	62.24
1539	ND1	HIS	359	-6.346	71.660	24.966	1.00	62.58
1540	CE1	HIS	359	-6.765	72.301	26.044	1.00	62.96
1541	NE2	HIS	359	-6.969	73.567	25.724	1.00	63.13
1542	C	HIS	359	-7.063	69.921	22.414	1.00	55.70
1543	O	HIS	359	-6.272	69.036	22.090	1.00	57.09
1544	N	ARG	360	-8.031	69.723	23.300	1.00	54.98
1545	CA	ARG	360	-8.218	68.421	23.915	1.00	53.99
1546	CB	ARG	360	-9.198	68.536	25.079	1.00	56.79
1547	CG	ARG	360	-9.424	67.240	25.835	1.00	59.14
1548	CD	ARG	360	-9.634	67.527	27.314	1.00	63.35
1549	NE	ARG	360	-8.522	68.304	27.859	1.00	65.64
1550	CZ	ARG	360	-8.227	68.404	29.152	1.00	66.48
1551	NH1	ARG	360	-8.963	67.769	30.059	1.00	66.32
1552	NH2	ARG	360	-7.188	69.137	29.533	1.00	66.19
1553	C	ARG	360	-8.760	67.457	22.855	1.00	53.28
1554	O	ARG	360	-8.353	66.294	22.780	1.00	52.91
1555	N	VAL	361	-9.671	67.956	22.026	1.00	51.42
1556	CA	VAL	361	-10.268	67.154	20.964	1.00	51.17
1557	CB	VAL	361	-11.391	67.934	20.227	1.00	51.88
1558	CG1	VAL	361	-11.995	67.067	19.147	1.00	50.23
1559	CG2	VAL	361	-12.457	68.388	21.208	1.00	51.85
1560	C	VAL	361	-9.218	66.768	19.925	1.00	50.25
1561	O	VAL	361	-8.958	65.591	19.677	1.00	48.98
1562	N	VAL	362	-8.627	67.784	19.312	1.00	49.39
1563	CA	VAL	362	-7.625	67.578	18.291	1.00	49.88
1564	CB	VAL	362	-7.010	68.924	17.873	1.00	51.37
1565	CG1	VAL	362	-6.369	69.582	19.075	1.00	54.78
1566	CG2	VAL	362	-5.990	68.722	16.765	1.00	52.24
1567	C	VAL	362	-6.534	66.629	18.782	1.00	48.98
1568	O	VAL	362	-6.089	65.755	18.045	1.00	48.59
1569	N	ASP	363	-6.110	66.790	20.030	1.00	48.62
1570	CA	ASP	363	-5.073	65.922	20.585	1.00	48.25
1571	CB	ASP	363	-4.779	66.279	22.039	1.00	47.89
1572	CG	ASP	363	-3.854	65.281	22.699	1.00	47.60
1573	OD1	ASP	363	-2.784	64.994	22.127	1.00	50.37
1574	OD2	ASP	363	-4.189	64.780	23.790	1.00	48.67
1575	C	ASP	363	-5.469	64.459	20.525	1.00	47.34
1576	O	ASP	363	-4.712	63.612	20.055	1.00	47.87
1577	N	GLN	364	-6.663	64.166	21.016	1.00	47.46
1578	CA	GLN	364	-7.164	62.804	21.028	1.00	47.48
1579	CB	GLN	364	-8.479	62.781	21.793	1.00	50.95
1580	CG	GLN	364	-9.487	61.782	21.312	1.00	55.14
1581	CD	GLN	364	-10.883	62.285	21.570	1.00	58.47
1582	OE1	GLN	364	-11.286	62.469	22.729	1.00	57.26
1583	NE2	GLN	364	-11.630	62.541	20.490	1.00	59.73
1584	C	GLN	364	-7.326	62.272	19.609	1.00	45.45
1585	O	GLN	364	-7.074	61.095	19.347	1.00	44.97
1586	N	LEU	365	-7.753	63.142	18.696	1.00	44.47
1587	CA	LEU	365	-7.903	62.771	17.291	1.00	42.18
1588	CB	LEU	365	-8.400	63.955	16.468	1.00	42.12

1589	CG	LEU	365	-9.896	64.194	16.327	1.00	42.56
1590	CD1	LEU	365	-10.124	65.467	15.524	1.00	42.11
1591	CD2	LEU	365	-10.533	63.004	15.630	1.00	42.65
1592	C	LEU	365	-6.541	62.357	16.751	1.00	41.43
1593	O	LEU	365	-6.427	61.370	16.027	1.00	41.80
1594	N	GLN	366	-5.512	63.124	17.103	1.00	40.35
1595	CA	GLN	366	-4.154	62.836	16.648	1.00	40.76
1596	CB	GLN	366	-3.187	63.940	17.082	1.00	40.50
1597	CG	GLN	366	-1.821	63.825	16.441	1.00	42.24
1598	CD	GLN	366	-0.803	64.768	17.049	1.00	44.24
1599	OE1	GLN	366	-0.529	64.712	18.249	1.00	45.86
1600	NE2	GLN	366	-0.232	65.638	16.224	1.00	44.85
1601	C	GLN	366	-3.668	61.500	17.193	1.00	40.73
1602	O	GLN	366	-2.907	60.799	16.532	1.00	40.82
1603	N	GLU	367	-4.103	61.145	18.395	1.00	40.49
1604	CA	GLU	367	-3.687	59.879	18.975	1.00	42.17
1605	CB	GLU	367	-4.076	59.816	20.448	1.00	45.45
1606	CG	GLU	367	-3.768	58.479	21.109	1.00	49.28
1607	CD	GLU	367	-3.734	58.579	22.624	1.00	53.07
1608	OE1	GLU	367	-3.496	57.541	23.288	1.00	53.21
1609	OE2	GLU	367	-3.936	59.700	23.149	1.00	55.05
1610	C	GLU	367	-4.310	58.709	18.229	1.00	41.49
1611	O	GLU	367	-3.672	57.676	18.021	1.00	40.03
1612	N	GLN	368	-5.564	58.875	17.831	1.00	41.18
1613	CA	GLN	368	-6.259	57.828	17.106	1.00	42.23
1614	CB	GLN	368	-7.715	58.225	16.867	1.00	44.56
1615	CG	GLN	368	-8.555	58.311	18.127	1.00	47.78
1616	CD	GLN	368	-9.944	58.864	17.854	1.00	51.26
1617	OE1	GLN	368	-10.625	58.430	16.924	1.00	52.63
1618	NE2	GLN	368	-10.373	59.824	18.671	1.00	53.12
1619	C	GLN	368	-5.562	57.581	15.778	1.00	41.50
1620	O	GLN	368	-5.507	56.444	15.297	1.00	40.66
1621	N	PHE	369	-5.032	58.649	15.186	1.00	40.87
1622	CA	PHE	369	-4.322	58.535	13.911	1.00	40.34
1623	CB	PHE	369	-4.056	59.920	13.308	1.00	37.34
1624	CG	PHE	369	-5.252	60.525	12.649	1.00	36.28
1625	CD1	PHE	369	-5.688	61.794	12.999	1.00	35.79
1626	CD2	PHE	369	-5.963	59.815	11.690	1.00	35.43
1627	CE1	PHE	369	-6.821	62.346	12.408	1.00	33.60
1628	CE2	PHE	369	-7.090	60.359	11.096	1.00	33.62
1629	CZ	PHE	369	-7.521	61.628	11.458	1.00	33.05
1630	C	PHE	369	-3.005	57.795	14.101	1.00	40.40
1631	O	PHE	369	-2.701	56.856	13.357	1.00	40.80
1632	N	SER	370	-2.230	58.212	15.100	1.00	39.04
1633	CA	SER	370	-0.953	57.570	15.372	1.00	39.43
1634	CB	SER	370	-0.265	58.222	16.572	1.00	40.15
1635	C	SER	370	-1.177	56.091	15.656	1.00	39.61
1636	O	SER	370	-0.441	55.239	15.161	1.00	38.71
1637	N	ILE	371	-2.203	55.790	16.449	1.00	39.37
1638	CA	ILE	371	-2.505	54.408	16.785	1.00	38.84
1639	CB	ILE	371	-3.684	54.313	17.798	1.00	39.62
1640	CG2	ILE	371	-4.119	52.860	17.987	1.00	38.85
1641	CG1	ILE	371	-3.247	54.882	19.145	1.00	39.70
1642	CD1	ILE	371	-4.248	54.671	20.251	1.00	39.78
1643	C	ILE	371	-2.831	53.617	15.523	1.00	38.45

1644	O	ILE	371	-2.393	52.475	15.377	1.00	38.75
1645	N	THR	372	-3.588	54.225	14.611	1.00	37.56
1646	CA	THR	372	-3.958	53.569	13.356	1.00	35.75
1647	CB	THR	372	-4.934	54.460	12.526	1.00	36.83
1648	OG1	THR	372	-6.171	54.614	13.237	1.00	36.76
1649	CG2	THR	372	-5.223	53.840	11.174	1.00	35.11
1650	C	THR	372	-2.695	53.278	12.529	1.00	34.48
1651	O	THR	372	-2.484	52.149	12.076	1.00	30.92
1652	N	LEU	373	-1.853	54.295	12.350	1.00	33.34
1653	CA	LEU	373	-0.621	54.128	11.583	1.00	34.64
1654	CB	LEU	373	0.177	55.431	11.549	1.00	33.60
1655	CG	LEU	373	1.595	55.316	10.981	1.00	32.50
1656	CD1	LEU	373	1.529	54.939	9.516	1.00	33.44
1657	CD2	LEU	373	2.329	56.626	11.148	1.00	33.06
1658	C	LEU	373	0.238	53.038	12.204	1.00	36.43
1659	O	LEU	373	0.760	52.158	11.513	1.00	35.78
1660	N	LYS	374	0.394	53.118	13.520	1.00	38.32
1661	CA	LYS	374	1.181	52.141	14.251	1.00	37.98
1662	CB	LYS	374	1.088	52.409	15.755	1.00	37.39
1663	CG	LYS	374	1.961	51.494	16.588	1.00	38.70
1664	CD	LYS	374	1.896	51.816	18.069	1.00	37.07
1665	CE	LYS	374	2.715	50.813	18.867	1.00	37.83
1666	NZ	LYS	374	2.395	50.837	20.323	1.00	38.03
1667	C	LYS	374	0.618	50.765	13.924	1.00	38.53
1668	O	LYS	374	1.358	49.840	13.588	1.00	39.28
1669	N	SER	375	-0.701	50.641	14.010	1.00	38.28
1670	CA	SER	375	-1.352	49.377	13.717	1.00	39.29
1671	CB	SER	375	-2.871	49.509	13.860	1.00	40.04
1672	OG	SER	375	-3.265	49.729	15.203	1.00	40.99
1673	C	SER	375	-1.009	48.943	12.298	1.00	40.53
1674	O	SER	375	-0.423	47.876	12.089	1.00	40.87
1675	N	TYR	376	-1.367	49.784	11.328	1.00	40.51
1676	CA	TYR	376	-1.120	49.494	9.919	1.00	39.81
1677	CB	TYR	376	-1.319	50.760	9.065	1.00	39.08
1678	CG	TYR	376	-1.053	50.554	7.583	1.00	38.38
1679	CD1	TYR	376	0.253	50.516	7.077	1.00	37.85
1680	CE1	TYR	376	0.494	50.234	5.729	1.00	37.34
1681	CD2	TYR	376	-2.104	50.319	6.698	1.00	36.45
1682	CE2	TYR	376	-1.867	50.039	5.350	1.00	34.80
1683	CZ	TYR	376	-0.577	49.992	4.877	1.00	35.66
1684	OH	TYR	376	-0.361	49.664	3.561	1.00	35.83
1685	C	TYR	376	0.271	48.926	9.678	1.00	40.75
1686	O	TYR	376	0.427	47.956	8.932	1.00	41.17
1687	N	ILE	377	1.278	49.527	10.305	1.00	40.95
1688	CA	ILE	377	2.651	49.073	10.141	1.00	41.92
1689	CB	ILE	377	3.644	50.038	10.821	1.00	40.33
1690	CG2	ILE	377	5.056	49.494	10.708	1.00	38.79
1691	CG1	ILE	377	3.536	51.426	10.177	1.00	37.73
1692	CD1	ILE	377	4.431	52.471	10.788	1.00	33.49
1693	C	ILE	377	2.837	47.676	10.717	1.00	45.20
1694	O	ILE	377	3.472	46.819	10.101	1.00	45.72
1695	N	GLU	378	2.272	47.435	11.894	1.00	48.26
1696	CA	GLU	378	2.405	46.126	12.515	1.00	51.38
1697	CB	GLU	378	1.995	46.201	13.994	1.00	52.80
1698	CG	GLU	378	3.193	46.321	14.948	1.00	56.56

1699	CD	GLU	378	2.885	47.088	16.232	1.00	58.91
1700	OE1	GLU	378	1.850	46.787	16.876	1.00	58.53
1701	OE2	GLU	378	3.693	47.985	16.593	1.00	58.27
1702	C	GLU	378	1.601	45.061	11.771	1.00	52.59
1703	O	GLU	378	1.940	43.878	11.814	1.00	53.23
1704	N	CYS	379	0.557	45.486	11.065	1.00	53.40
1705	CA	CYS	379	-0.292	44.563	10.318	1.00	54.39
1706	CB	CYS	379	-1.747	45.030	10.366	1.00	55.43
1707	SG	CYS	379	-2.537	44.887	11.972	1.00	59.84
1708	C	CYS	379	0.088	44.350	8.855	1.00	54.73
1709	O	CYS	379	-0.467	43.469	8.202	1.00	54.22
1710	N	ASN	380	1.017	45.141	8.327	1.00	55.14
1711	CA	ASN	380	1.382	44.979	6.924	1.00	56.35
1712	CB	ASN	380	0.678	46.052	6.089	1.00	57.43
1713	CG	ASN	380	-0.837	45.995	6.229	1.00	58.94
1714	OD1	ASN	380	-1.387	46.244	7.303	1.00	58.98
1715	ND2	ASN	380	-1.517	45.656	5.143	1.00	60.02
1716	C	ASN	380	2.876	44.970	6.602	1.00	56.88
1717	O	ASN	380	3.256	44.788	5.442	1.00	57.40
1718	N	ARG	381	3.721	45.154	7.617	1.00	56.03
1719	CA	ARG	381	5.172	45.158	7.418	1.00	55.06
1720	CB	ARG	381	5.699	46.593	7.432	1.00	52.20
1721	CG	ARG	381	4.758	47.588	6.799	1.00	50.70
1722	CD	ARG	381	5.437	48.916	6.573	1.00	51.33
1723	NE	ARG	381	6.422	48.818	5.506	1.00	50.60
1724	CZ	ARG	381	6.341	49.484	4.362	1.00	51.54
1725	NH1	ARG	381	5.319	50.303	4.144	1.00	51.64
1726	NH2	ARG	381	7.272	49.321	3.431	1.00	52.06
1727	C	ARG	381	5.860	44.336	8.516	1.00	55.83
1728	O	ARG	381	6.418	44.884	9.467	1.00	56.00
1729	N	PRO	382	5.840	43.004	8.384	1.00	56.44
1730	CD	PRO	382	5.165	42.250	7.311	1.00	56.40
1731	CA	PRO	382	6.446	42.088	9.353	1.00	57.48
1732	CB	PRO	382	5.717	40.783	9.068	1.00	57.30
1733	CG	PRO	382	5.624	40.820	7.573	1.00	57.08
1734	C	PRO	382	7.965	41.916	9.269	1.00	57.78
1735	O	PRO	382	8.581	41.364	10.184	1.00	58.22
1736	N	GLN	383	8.567	42.383	8.182	1.00	58.24
1737	CA	GLN	383	10.006	42.226	7.996	1.00	58.87
1738	CB	GLN	383	10.368	42.396	6.512	1.00	59.15
1739	CG	GLN	383	9.530	43.406	5.729	1.00	61.08
1740	CD	GLN	383	8.121	42.910	5.412	1.00	62.24
1741	OE1	GLN	383	7.926	41.751	5.048	1.00	63.60
1742	NE2	GLN	383	7.137	43.798	5.526	1.00	61.99
1743	C	GLN	383	10.926	43.087	8.870	1.00	58.42
1744	O	GLN	383	10.606	44.223	9.223	1.00	57.85
1745	N	PRO	384	12.093	42.529	9.231	1.00	58.24
1746	CD	PRO	384	12.475	41.163	8.823	1.00	57.66
1747	CA	PRO	384	13.145	43.127	10.058	1.00	57.49
1748	CB	PRO	384	14.320	42.176	9.855	1.00	57.15
1749	CG	PRO	384	13.651	40.860	9.731	1.00	57.34
1750	C	PRO	384	13.523	44.561	9.719	1.00	57.41
1751	O	PRO	384	14.007	45.292	10.584	1.00	58.50
1752	N	ALA	385	13.319	44.963	8.468	1.00	56.58
1753	CA	ALA	385	13.677	46.319	8.045	1.00	54.85

1754	CB	ALA	385	13.955	46.351	6.543	1.00	54.05
1755	C	ALA	385	12.596	47.331	8.388	1.00	53.19
1756	O	ALA	385	12.812	48.545	8.299	1.00	53.47
1757	N	HIS	386	11.431	46.827	8.779	1.00	50.77
1758	CA	HIS	386	10.321	47.697	9.115	1.00	47.65
1759	CB	HIS	386	9.107	47.302	8.280	1.00	47.46
1760	CG	HIS	386	9.345	47.433	6.806	1.00	48.41
1761	CD2	HIS	386	9.924	48.422	6.084	1.00	48.77
1762	ND1	HIS	386	9.015	46.445	5.903	1.00	49.53
1763	CE1	HIS	386	9.384	46.816	4.689	1.00	48.73
1764	NE2	HIS	386	9.939	48.011	4.772	1.00	49.46
1765	C	HIS	386	10.041	47.677	10.604	1.00	45.69
1766	O	HIS	386	8.915	47.886	11.057	1.00	46.86
1767	N	ARG	387	11.096	47.416	11.367	1.00	42.07
1768	CA	ARG	387	11.007	47.426	12.811	1.00	38.42
1769	CB	ARG	387	12.070	46.508	13.428	1.00	37.70
1770	CG	ARG	387	11.741	45.022	13.423	1.00	35.02
1771	CD	ARG	387	12.840	44.208	14.123	1.00	33.86
1772	NE	ARG	387	13.058	44.643	15.503	1.00	33.71
1773	CZ	ARG	387	14.024	44.189	16.307	1.00	33.75
1774	NH1	ARG	387	14.885	43.270	15.886	1.00	33.65
1775	NH2	ARG	387	14.138	44.667	17.539	1.00	32.06
1776	C	ARG	387	11.293	48.883	13.185	1.00	37.16
1777	O	ARG	387	12.191	49.517	12.620	1.00	36.36
1778	N	PHE	388	10.521	49.418	14.121	1.00	34.47
1779	CA	PHE	388	10.706	50.790	14.562	1.00	33.54
1780	CB	PHE	388	12.173	51.046	14.905	1.00	33.05
1781	CG	PHE	388	12.770	50.031	15.831	1.00	32.89
1782	CD1	PHE	388	14.008	49.462	15.549	1.00	32.67
1783	CD2	PHE	388	12.096	49.623	16.974	1.00	33.98
1784	CE1	PHE	388	14.564	48.499	16.390	1.00	30.47
1785	CE2	PHE	388	12.648	48.655	17.822	1.00	30.89
1786	CZ	PHE	388	13.881	48.097	17.525	1.00	29.41
1787	C	PHE	388	10.259	51.791	13.497	1.00	34.54
1788	O	PHE	388	10.473	53.002	13.643	1.00	35.37
1789	N	LEU	389	9.641	51.302	12.424	1.00	33.70
1790	CA	LEU	389	9.179	52.204	11.372	1.00	31.34
1791	CB	LEU	389	8.396	51.439	10.305	1.00	30.45
1792	CG	LEU	389	7.993	52.280	9.087	1.00	30.09
1793	CD1	LEU	389	9.202	53.018	8.510	1.00	26.66
1794	CD2	LEU	389	7.390	51.374	8.047	1.00	29.32
1795	C	LEU	389	8.295	53.286	11.982	1.00	30.08
1796	O	LEU	389	8.539	54.485	11.815	1.00	29.89
1797	N	PHE	390	7.279	52.854	12.715	1.00	28.06
1798	CA	PHE	390	6.369	53.789	13.349	1.00	29.26
1799	CB	PHE	390	5.417	53.040	14.278	1.00	26.13
1800	CG	PHE	390	4.582	53.940	15.130	1.00	24.65
1801	CD1	PHE	390	3.691	54.829	14.556	1.00	25.32
1802	CD2	PHE	390	4.704	53.913	16.507	1.00	24.64
1803	CE1	PHE	390	2.937	55.679	15.344	1.00	25.64
1804	CE2	PHE	390	3.949	54.760	17.304	1.00	24.77
1805	CZ	PHE	390	3.069	55.642	16.725	1.00	25.53
1806	C	PHE	390	7.082	54.911	14.122	1.00	30.59
1807	O	PHE	390	6.701	56.075	14.022	1.00	31.26
1808	N	LEU	391	8.111	54.571	14.891	1.00	31.39

1809	CA	LEU	391	8.827	55.587	15.657	1.00	32.01
1810	CB	LEU	391	9.703	54.922	16.724	1.00	32.34
1811	CG	LEU	391	9.150	54.909	18.154	1.00	32.42
1812	CD1	LEU	391	7.636	54.780	18.159	1.00	34.24
1813	CD2	LEU	391	9.786	53.757	18.910	1.00	32.71
1814	C	LEU	391	9.664	56.504	14.768	1.00	31.89
1815	O	LEU	391	9.693	57.718	14.987	1.00	30.53
1816	N	LYS	392	10.333	55.928	13.767	1.00	32.59
1817	CA	LYS	392	11.149	56.718	12.836	1.00	32.98
1818	CB	LYS	392	11.731	55.843	11.728	1.00	33.11
1819	CG	LYS	392	12.973	55.048	12.113	1.00	36.43
1820	CD	LYS	392	13.498	54.271	10.910	1.00	35.90
1821	CE	LYS	392	14.443	53.148	11.303	1.00	36.51
1822	NZ	LYS	392	14.690	52.220	10.149	1.00	37.34
1823	C	LYS	392	10.279	57.787	12.194	1.00	33.01
1824	O	LYS	392	10.685	58.939	12.053	1.00	32.82
1825	N	ILE	393	9.078	57.378	11.799	1.00	32.99
1826	CA	ILE	393	8.109	58.268	11.179	1.00	30.21
1827	CB	ILE	393	6.856	57.468	10.735	1.00	29.88
1828	CG2	ILE	393	5.726	58.408	10.312	1.00	28.35
1829	CG1	ILE	393	7.240	56.522	9.595	1.00	28.20
1830	CD1	ILE	393	6.113	55.634	9.133	1.00	27.20
1831	C	ILE	393	7.708	59.352	12.170	1.00	30.21
1832	O	ILE	393	7.653	60.536	11.831	1.00	30.54
1833	N	MET	394	7.439	58.956	13.406	1.00	30.84
1834	CA	MET	394	7.036	59.939	14.395	1.00	31.81
1835	CB	MET	394	6.602	59.256	15.701	1.00	30.69
1836	CG	MET	394	5.294	58.468	15.584	1.00	29.74
1837	SD	MET	394	3.914	59.394	14.826	1.00	28.58
1838	CE	MET	394	3.507	60.527	16.171	1.00	27.25
1839	C	MET	394	8.170	60.921	14.631	1.00	31.96
1840	O	MET	394	7.935	62.122	14.730	1.00	32.84
1841	N	GLN	395	9.399	60.417	14.705	1.00	32.20
1842	CA	GLN	395	10.554	61.290	14.900	1.00	32.10
1843	CB	GLN	395	11.833	60.464	15.059	1.00	29.90
1844	C	GLN	395	10.667	62.167	13.662	1.00	31.67
1845	O	GLN	395	10.924	63.364	13.741	1.00	30.56
1846	N	MET	396	10.455	61.538	12.512	1.00	33.98
1847	CA	MET	396	10.520	62.195	11.216	1.00	34.15
1848	CB	MET	396	10.154	61.203	10.122	1.00	37.45
1849	CG	MET	396	10.823	61.479	8.802	1.00	44.01
1850	SD	MET	396	12.552	61.051	8.909	1.00	47.13
1851	CE	MET	396	12.415	59.258	8.810	1.00	47.24
1852	C	MET	396	9.565	63.378	11.158	1.00	33.39
1853	O	MET	396	9.899	64.434	10.602	1.00	32.17
1854	N	LEU	397	8.374	63.196	11.725	1.00	31.88
1855	CA	LEU	397	7.375	64.259	11.746	1.00	32.83
1856	CB	LEU	397	5.988	63.692	12.063	1.00	32.09
1857	CG	LEU	397	5.276	63.008	10.894	1.00	33.43
1858	CD1	LEU	397	3.896	62.525	11.320	1.00	31.78
1859	CD2	LEU	397	5.162	63.992	9.740	1.00	32.01
1860	C	LEU	397	7.727	65.366	12.744	1.00	34.19
1861	O	LEU	397	7.358	66.530	12.546	1.00	33.53
1862	N	THR	398	8.435	65.011	13.815	1.00	35.42
1863	CA	THR	398	8.825	66.011	14.808	1.00	37.07

1864	CB	THR	398	9.414	65.365	16.091	1.00	37.19
1865	OG1	THR	398	8.452	64.469	16.663	1.00	39.03
1866	CG2	THR	398	9.748	66.438	17.122	1.00	36.51
1867	C	THR	398	9.880	66.893	14.161	1.00	36.99
1868	O	THR	398	9.934	68.102	14.388	1.00	36.45
1869	N	GLU	399	10.715	66.270	13.342	1.00	38.17
1870	CA	GLU	399	11.756	66.994	12.639	1.00	40.75
1871	CB	GLU	399	12.619	66.026	11.826	1.00	42.78
1872	CG	GLU	399	14.080	66.023	12.220	1.00	47.62
1873	CD	GLU	399	14.871	67.130	11.554	1.00	51.21
1874	OE1	GLU	399	15.076	67.053	10.323	1.00	53.88
1875	OE2	GLU	399	15.289	68.078	12.253	1.00	52.92
1876	C	GLU	399	11.082	67.982	11.708	1.00	40.31
1877	O	GLU	399	11.429	69.165	11.688	1.00	41.13
1878	N	LEU	400	10.100	67.495	10.954	1.00	38.44
1879	CA	LEU	400	9.395	68.338	10.004	1.00	38.33
1880	CB	LEU	400	8.248	67.560	9.363	1.00	39.48
1881	CG	LEU	400	7.728	68.144	8.044	1.00	39.48
1882	CD1	LEU	400	8.869	68.276	7.049	1.00	38.58
1883	CD2	LEU	400	6.649	67.245	7.484	1.00	39.89
1884	C	LEU	400	8.876	69.632	10.633	1.00	38.20
1885	O	LEU	400	8.938	70.689	10.017	1.00	38.26
1886	N	ARG	401	8.365	69.542	11.855	1.00	38.42
1887	CA	ARG	401	7.859	70.703	12.580	1.00	37.96
1888	CB	ARG	401	7.350	70.269	13.950	1.00	41.21
1889	CG	ARG	401	5.901	69.876	13.977	1.00	47.61
1890	CD	ARG	401	5.048	71.117	14.067	1.00	53.75
1891	NE	ARG	401	5.287	71.817	15.326	1.00	56.78
1892	CZ	ARG	401	4.855	71.389	16.508	1.00	58.74
1893	NH1	ARG	401	4.152	70.264	16.591	1.00	58.66
1894	NH2	ARG	401	5.145	72.076	17.608	1.00	59.75
1895	C	ARG	401	8.965	71.737	12.760	1.00	37.52
1896	O	ARG	401	8.743	72.934	12.583	1.00	35.32
1897	N	SER	402	10.154	71.268	13.129	1.00	37.43
1898	CA	SER	402	11.296	72.153	13.320	1.00	38.54
1899	CB	SER	402	12.507	71.371	13.821	1.00	38.32
1900	OG	SER	402	12.469	71.267	15.226	1.00	41.02
1901	C	SER	402	11.664	72.881	12.034	1.00	38.72
1902	O	SER	402	11.955	74.077	12.057	1.00	38.57
1903	N	LEU	403	11.670	72.159	10.918	1.00	38.23
1904	CA	LEU	403	11.989	72.782	9.647	1.00	40.43
1905	CB	LEU	403	12.001	71.743	8.529	1.00	41.31
1906	CG	LEU	403	12.749	70.467	8.913	1.00	43.89
1907	CD1	LEU	403	12.890	69.554	7.695	1.00	44.38
1908	CD2	LEU	403	14.119	70.838	9.470	1.00	45.32
1909	C	LEU	403	10.921	73.835	9.370	1.00	41.14
1910	O	LEU	403	11.240	74.960	8.991	1.00	42.61
1911	N	ASN	404	9.657	73.462	9.576	1.00	40.01
1912	CA	ASN	404	8.524	74.359	9.361	1.00	39.37
1913	CB	ASN	404	7.249	73.720	9.930	1.00	39.51
1914	CG	ASN	404	5.978	74.477	9.555	1.00	41.84
1915	OD1	ASN	404	5.902	75.121	8.508	1.00	41.75
1916	ND2	ASN	404	4.962	74.381	10.408	1.00	43.88
1917	C	ASN	404	8.806	75.706	10.033	1.00	39.54
1918	O	ASN	404	8.863	76.739	9.371	1.00	39.28

1919	N	GLU	405	8.999	75.696	11.345	1.00	39.72
1920	CA	GLU	405	9.296	76.931	12.058	1.00	40.63
1921	CB	GLU	405	9.583	76.635	13.524	1.00	38.75
1922	C	GLU	405	10.515	77.625	11.440	1.00	42.25
1923	O	GLU	405	10.484	78.833	11.159	1.00	41.95
1924	N	GLN	406	11.582	76.852	11.243	1.00	42.88
1925	CA	GLN	406	12.819	77.360	10.674	1.00	44.40
1926	CB	GLN	406	13.760	76.181	10.380	1.00	45.75
1927	CG	GLN	406	15.216	76.523	10.088	1.00	48.90
1928	CD	GLN	406	15.527	76.569	8.597	1.00	52.34
1929	OE1	GLN	406	15.573	77.644	7.988	1.00	50.67
1930	NE2	GLN	406	15.736	75.391	7.999	1.00	52.63
1931	C	GLN	406	12.488	78.098	9.389	1.00	45.52
1932	O	GLN	406	12.996	79.189	9.118	1.00	44.82
1933	N	HIS	407	11.587	77.492	8.624	1.00	47.67
1934	CA	HIS	407	11.215	78.033	7.329	1.00	49.29
1935	CB	HIS	407	10.624	76.948	6.410	1.00	53.97
1936	CG	HIS	407	11.640	76.530	5.430	1.00	60.70
1937	CD2	HIS	407	12.646	75.631	5.551	1.00	63.31
1938	ND1	HIS	407	11.976	77.353	4.373	1.00	63.04
1939	CE1	HIS	407	13.162	77.004	3.915	1.00	65.48
1940	NE2	HIS	407	13.599	75.972	4.620	1.00	64.91
1941	C	HIS	407	10.365	79.240	7.289	1.00	48.62
1942	O	HIS	407	10.628	80.141	6.505	1.00	47.86
1943	N	THR	408	9.349	79.255	8.139	1.00	47.55
1944	CA	THR	408	8.478	80.385	8.217	1.00	47.16
1945	CB	THR	408	7.513	80.188	9.383	1.00	48.27
1946	OG1	THR	408	6.649	79.086	9.085	1.00	48.15
1947	CG2	THR	408	6.689	81.432	9.617	1.00	48.27
1948	C	THR	408	9.381	81.595	8.426	1.00	46.37
1949	O	THR	408	9.247	82.603	7.730	1.00	45.00
1950	N	GLN	409	10.321	81.487	9.365	1.00	46.43
1951	CA	GLN	409	11.255	82.572	9.659	1.00	46.48
1952	CB	GLN	409	12.211	82.152	10.753	1.00	46.66
1953	CG	GLN	409	11.556	81.745	12.041	1.00	49.02
1954	CD	GLN	409	12.495	80.910	12.886	1.00	50.55
1955	OE1	GLN	409	13.723	81.057	12.800	1.00	50.20
1956	NE2	GLN	409	11.930	80.028	13.708	1.00	50.66
1957	C	GLN	409	12.076	82.913	8.412	1.00	47.24
1958	O	GLN	409	12.133	84.069	7.986	1.00	47.81
1959	N	ARG	410	12.725	81.903	7.837	1.00	46.78
1960	CA	ARG	410	13.533	82.116	6.650	1.00	46.27
1961	CB	ARG	410	13.965	80.773	6.066	1.00	48.14
1962	CG	ARG	410	15.150	80.884	5.132	1.00	53.28
1963	CD	ARG	410	15.440	79.601	4.345	1.00	57.80
1964	NE	ARG	410	15.767	78.445	5.179	1.00	61.58
1965	CZ	ARG	410	16.418	77.367	4.736	1.00	63.41
1966	NH1	ARG	410	16.818	77.300	3.468	1.00	63.69
1967	NH2	ARG	410	16.654	76.348	5.555	1.00	63.52
1968	C	ARG	410	12.684	82.883	5.637	1.00	45.43
1969	O	ARG	410	13.084	83.939	5.136	1.00	45.56
1970	N	LEU	411	11.499	82.345	5.361	1.00	43.43
1971	CA	LEU	411	10.567	82.940	4.414	1.00	41.11
1972	CB	LEU	411	9.328	82.051	4.303	1.00	40.73
1973	CG	LEU	411	8.136	82.514	3.459	1.00	43.60

1974	CD1	LEU	411	7.416	81.299	2.864	1.00	45.74
1975	CD2	LEU	411	7.179	83.341	4.320	1.00	43.85
1976	C	LEU	411	10.170	84.376	4.768	1.00	40.67
1977	O	LEU	411	10.000	85.215	3.882	1.00	39.23
1978	N	LEU	412	10.028	84.672	6.057	1.00	40.61
1979	CA	LEU	412	9.643	86.021	6.464	1.00	40.71
1980	CB	LEU	412	9.147	86.030	7.912	1.00	39.50
1981	CG	LEU	412	7.770	85.389	8.121	1.00	39.12
1982	CD1	LEU	412	7.278	85.697	9.524	1.00	37.62
1983	CD2	LEU	412	6.779	85.921	7.089	1.00	37.52
1984	C	LEU	412	10.762	87.044	6.286	1.00	41.31
1985	O	LEU	412	10.502	88.185	5.904	1.00	39.96
1986	N	ARG	413	12.000	86.637	6.564	1.00	42.35
1987	CA	ARG	413	13.144	87.525	6.400	1.00	42.82
1988	CB	ARG	413	14.439	86.790	6.732	1.00	43.85
1989	CG	ARG	413	14.847	86.877	8.191	1.00	44.48
1990	CD	ARG	413	16.071	86.030	8.447	1.00	44.45
1991	NE	ARG	413	15.718	84.664	8.815	1.00	47.14
1992	CZ	ARG	413	16.489	83.609	8.571	1.00	49.99
1993	NH1	ARG	413	17.650	83.770	7.948	1.00	52.20
1994	NH2	ARG	413	16.116	82.397	8.964	1.00	50.67
1995	C	ARG	413	13.173	87.995	4.952	1.00	43.85
1996	O	ARG	413	13.314	89.186	4.671	1.00	43.59
1997	N	ILE	414	13.031	87.045	4.037	1.00	44.67
1998	CA	ILE	414	13.015	87.346	2.613	1.00	45.43
1999	CB	ILE	414	12.854	86.046	1.781	1.00	45.64
2000	CG2	ILE	414	12.563	86.378	0.318	1.00	43.61
2001	CG1	ILE	414	14.119	85.191	1.926	1.00	45.58
2002	CD1	ILE	414	14.116	83.906	1.107	1.00	45.47
2003	C	ILE	414	11.873	88.303	2.279	1.00	46.03
2004	O	ILE	414	12.054	89.253	1.516	1.00	47.40
2005	N	GLN	415	10.702	88.051	2.859	1.00	46.69
2006	CA	GLN	415	9.520	88.883	2.620	1.00	47.57
2007	CB	GLN	415	8.301	88.289	3.326	1.00	45.38
2008	CG	GLN	415	7.003	89.031	3.081	1.00	42.31
2009	CD	GLN	415	6.495	88.845	1.673	1.00	44.02
2010	OE1	GLN	415	7.043	88.052	0.906	1.00	45.05
2011	NE2	GLN	415	5.434	89.567	1.321	1.00	42.21
2012	C	GLN	415	9.739	90.300	3.127	1.00	49.59
2013	O	GLN	415	9.307	91.270	2.505	1.00	49.56
2014	N	ASP	416	10.411	90.409	4.267	1.00	51.59
2015	CA	ASP	416	10.682	91.703	4.865	1.00	53.89
2016	CB	ASP	416	11.361	91.517	6.229	1.00	56.53
2017	CG	ASP	416	11.762	92.838	6.870	1.00	59.62
2018	OD1	ASP	416	10.886	93.719	7.031	1.00	60.06
2019	OD2	ASP	416	12.957	92.992	7.215	1.00	61.63
2020	C	ASP	416	11.527	92.610	3.968	1.00	53.82
2021	O	ASP	416	11.501	93.829	4.126	1.00	54.73
2022	N	ILE	417	12.271	92.037	3.026	1.00	53.12
2023	CA	ILE	417	13.086	92.866	2.142	1.00	53.22
2024	CB	ILE	417	14.610	92.597	2.322	1.00	53.89
2025	CG2	ILE	417	15.080	93.156	3.656	1.00	54.20
2026	CG1	ILE	417	14.916	91.101	2.224	1.00	54.07
2027	CD1	ILE	417	14.828	90.535	0.828	1.00	54.43
2028	C	ILE	417	12.729	92.706	0.673	1.00	53.35

2029	O	ILE	417	13.124	93.522	-0.162	1.00	53.35
2030	N	HIS	418	11.973	91.661	0.357	1.00	52.73
2031	CA	HIS	418	11.583	91.421	-1.025	1.00	52.17
2032	CB	HIS	418	12.631	90.533	-1.704	1.00	52.92
2033	CG	HIS	418	12.588	90.573	-3.201	1.00	54.38
2034	CD2	HIS	418	13.399	91.184	-4.097	1.00	53.68
2035	ND1	HIS	418	11.624	89.924	-3.941	1.00	55.24
2036	CE1	HIS	418	11.843	90.131	-5.227	1.00	53.97
2037	NE2	HIS	418	12.914	90.892	-5.349	1.00	54.12
2038	C	HIS	418	10.217	90.754	-1.070	1.00	51.33
2039	O	HIS	418	10.123	89.532	-1.136	1.00	53.28
2040	N	PRO	419	9.136	91.552	-1.028	1.00	49.45
2041	CD	PRO	419	9.090	93.015	-1.173	1.00	48.47
2042	CA	PRO	419	7.785	90.985	-1.067	1.00	47.63
2043	CB	PRO	419	6.895	92.225	-1.134	1.00	47.50
2044	CG	PRO	419	7.753	93.218	-1.843	1.00	47.69
2045	C	PRO	419	7.625	90.084	-2.284	1.00	46.08
2046	O	PRO	419	8.133	90.401	-3.355	1.00	46.98
2047	N	PHE	420	6.932	88.961	-2.120	1.00	44.39
2048	CA	PHE	420	6.730	88.021	-3.227	1.00	43.32
2049	CB	PHE	420	8.023	87.253	-3.508	1.00	42.27
2050	CG	PHE	420	8.301	86.182	-2.502	1.00	41.66
2051	CD1	PHE	420	7.779	84.906	-2.665	1.00	41.66
2052	CD2	PHE	420	9.003	86.473	-1.340	1.00	41.36
2053	CE1	PHE	420	7.946	83.942	-1.679	1.00	41.75
2054	CE2	PHE	420	9.172	85.513	-0.351	1.00	40.46
2055	CZ	PHE	420	8.641	84.251	-0.522	1.00	40.32
2056	C	PHE	420	5.634	87.020	-2.865	1.00	41.58
2057	O	PHE	420	5.030	86.397	-3.735	1.00	40.37
2058	N	ALA	421	5.402	86.855	-1.569	1.00	39.94
2059	CA	ALA	421	4.383	85.940	-1.091	1.00	39.63
2060	CB	ALA	421	4.312	86.002	0.431	1.00	38.95
2061	C	ALA	421	3.015	86.268	-1.700	1.00	39.52
2062	O	ALA	421	2.658	87.439	-1.872	1.00	39.85
2063	N	THR	422	2.256	85.226	-2.028	1.00	37.98
2064	CA	THR	422	0.925	85.391	-2.610	1.00	35.41
2065	CB	THR	422	0.444	84.108	-3.328	1.00	35.08
2066	OG1	THR	422	0.384	83.027	-2.385	1.00	32.71
2067	CG2	THR	422	1.381	83.748	-4.479	1.00	35.60
2068	C	THR	422	-0.093	85.672	-1.519	1.00	34.77
2069	O	THR	422	0.198	85.516	-0.332	1.00	33.73
2070	N	PRO	423	-1.304	86.104	-1.912	1.00	34.23
2071	CD	PRO	423	-1.679	86.572	-3.263	1.00	34.14
2072	CA	PRO	423	-2.370	86.394	-0.955	1.00	31.86
2073	CB	PRO	423	-3.567	86.646	-1.856	1.00	31.16
2074	CG	PRO	423	-2.946	87.379	-2.995	1.00	33.24
2075	C	PRO	423	-2.581	85.200	-0.031	1.00	31.37
2076	O	PRO	423	-2.597	85.350	1.195	1.00	31.72
2077	N	LEU	424	-2.729	84.015	-0.627	1.00	30.26
2078	CA	LEU	424	-2.926	82.789	0.139	1.00	28.31
2079	CB	LEU	424	-3.014	81.580	-0.781	1.00	25.91
2080	CG	LEU	424	-3.373	80.305	-0.015	1.00	29.72
2081	CD1	LEU	424	-4.792	80.435	0.517	1.00	29.37
2082	CD2	LEU	424	-3.257	79.080	-0.919	1.00	31.09
2083	C	LEU	424	-1.764	82.599	1.100	1.00	29.68

2084	O	LEU	424	-1.967	82.373	2.292	1.00	31.25
2085	N	MET	425	-0.543	82.694	0.584	1.00	29.13
2086	CA	MET	425	0.643	82.551	1.420	1.00	29.90
2087	CB	MET	425	1.901	82.755	0.583	1.00	29.69
2088	CG	MET	425	2.213	81.577	-0.312	1.00	30.20
2089	SD	MET	425	3.537	81.924	-1.456	1.00	32.42
2090	CE	MET	425	4.944	82.048	-0.372	1.00	29.03
2091	C	MET	425	0.631	83.536	2.585	1.00	30.67
2092	O	MET	425	1.054	83.210	3.694	1.00	30.26
2093	N	GLN	426	0.151	84.745	2.324	1.00	31.96
2094	CA	GLN	426	0.067	85.777	3.349	1.00	33.65
2095	CB	GLN	426	-0.409	87.102	2.719	1.00	32.99
2096	CG	GLN	426	0.650	87.813	1.870	1.00	31.49
2097	CD	GLN	426	0.080	88.924	0.998	1.00	32.35
2098	OE1	GLN	426	-0.741	89.721	1.445	1.00	34.92
2099	NE2	GLN	426	0.527	88.987	-0.250	1.00	30.45
2100	C	GLN	426	-0.905	85.313	4.444	1.00	35.37
2101	O	GLN	426	-0.632	85.464	5.637	1.00	34.67
2102	N	GLU	427	-2.035	84.745	4.031	1.00	37.68
2103	CA	GLU	427	-3.040	84.255	4.969	1.00	40.84
2104	CB	GLU	427	-4.156	83.531	4.214	1.00	43.08
2105	CG	GLU	427	-5.136	84.421	3.482	1.00	47.76
2106	CD	GLU	427	-6.240	84.936	4.385	1.00	51.36
2107	OE1	GLU	427	-7.187	85.586	3.873	1.00	51.80
2108	OE2	GLU	427	-6.160	84.686	5.609	1.00	54.10
2109	C	GLU	427	-2.415	83.277	5.955	1.00	41.89
2110	O	GLU	427	-2.392	83.515	7.163	1.00	41.95
2111	N	LEU	428	-1.915	82.174	5.405	1.00	43.21
2112	CA	LEU	428	-1.293	81.091	6.160	1.00	44.72
2113	CB	LEU	428	-0.738	80.039	5.197	1.00	43.74
2114	CG	LEU	428	-1.751	79.231	4.385	1.00	44.22
2115	CD1	LEU	428	-1.074	78.630	3.159	1.00	44.72
2116	CD2	LEU	428	-2.359	78.153	5.267	1.00	42.69
2117	C	LEU	428	-0.185	81.541	7.091	1.00	46.27
2118	O	LEU	428	-0.101	81.090	8.231	1.00	47.11
2119	N	PHE	429	0.676	82.420	6.603	1.00	49.25
2120	CA	PHE	429	1.776	82.900	7.418	1.00	52.66
2121	CB	PHE	429	3.061	82.929	6.578	1.00	53.46
2122	CG	PHE	429	3.518	81.555	6.134	1.00	56.38
2123	CD1	PHE	429	2.769	80.814	5.219	1.00	57.39
2124	CD2	PHE	429	4.666	80.976	6.675	1.00	57.10
2125	CE1	PHE	429	3.151	79.512	4.851	1.00	58.24
2126	CE2	PHE	429	5.059	79.673	6.314	1.00	58.11
2127	CZ	PHE	429	4.298	78.942	5.402	1.00	58.06
2128	C	PHE	429	1.458	84.264	8.034	1.00	54.34
2129	O	PHE	429	0.288	84.637	8.181	1.00	53.18
2130	N	GLY	430	2.491	84.999	8.416	1.00	56.82
2131	CA	GLY	430	2.259	86.298	9.018	1.00	60.20
2132	C	GLY	430	1.286	87.162	8.233	1.00	61.66
2133	O	GLY	430	0.063	87.075	8.415	1.00	61.39
2134	N	ILE	431	1.851	87.989	7.355	1.00	61.49
2135	CA	ILE	431	1.115	88.927	6.513	1.00	61.54
2136	CB	ILE	431	1.624	88.874	5.068	1.00	61.84
2137	CG2	ILE	431	1.070	90.056	4.289	1.00	62.88
2138	CG1	ILE	431	3.154	88.926	5.044	1.00	62.50

2139	CD1	ILE	431	3.833	87.651	5.509	1.00	62.73
2140	C	ILE	431	-0.405	88.762	6.496	1.00	61.67
2141	O	ILE	431	-1.145	89.729	6.278	1.00	60.52
2142	O	HOH	1426	-1.082	69.659	-2.625	1.00	24.14
2143	O	HOH	1427	-10.671	76.436	6.561	1.00	39.57
2144	O	HOH	1428	-2.237	71.694	12.302	1.00	19.63
2145	O	HOH	1430	8.952	56.719	1.568	1.00	18.24
2146	O	HOH	1431	-0.779	59.892	24.059	1.00	26.54
2147	O	HOH	1432	22.789	67.648	-0.238	1.00	45.09
2148	O	HOH	1433	-11.256	70.329	23.204	1.00	39.97
2149	O	HOH	1434	-10.044	50.821	16.739	1.00	41.82
2150	O	HOH	1436	2.167	79.769	-17.224	1.00	30.96
2151	O	HOH	1437	16.991	84.149	-14.225	1.00	48.61
2152	O	HOH	1438	-11.704	56.433	-1.728	1.00	49.66
2153	O	HOH	1439	16.619	50.272	11.443	1.00	32.78
2154	O	HOH	1440	-0.731	59.018	-0.091	1.00	19.99
2155	O	HOH	1441	-7.575	88.256	3.592	1.00	32.60
2156	O	HOH	1442	-4.103	71.133	-5.988	1.00	43.57
2157	O	HOH	1443	3.986	90.037	-1.136	1.00	47.02
2158	O	HOH	1444	6.284	48.618	0.356	1.00	28.16
2159	O	HOH	1445	7.455	55.120	-1.050	1.00	40.82
2160	O	HOH	1446	9.754	82.836	-10.782	1.00	17.83
2161	O	HOH	1447	-5.744	62.064	-7.344	1.00	26.53
2162	O	HOH	1449	8.757	42.647	12.829	1.00	33.80
2163	O	HOH	1450	-6.308	70.707	-8.559	1.00	43.71
2164	O	HOH	1452	13.218	57.431	-5.501	1.00	44.44
2165	O	HOH	1453	0.854	52.302	-1.519	1.00	21.86
2166	O	HOH	1454	16.579	73.366	5.592	1.00	49.48
2167	O	HOH	1455	-0.111	81.068	-7.439	1.00	27.83
2168	O	HOH	1456	15.223	82.231	-6.302	1.00	46.60
2169	O	HOH	1458	-13.188	62.086	1.663	1.00	48.47
2170	O	HOH	1459	10.775	58.972	-13.157	1.00	53.96
2171	O	HOH	1461	6.902	50.457	13.830	1.00	39.37
2172	O	HOH	1462	17.366	88.658	-20.366	1.00	50.74
2173	O	HOH	1463	-3.131	64.501	-10.998	1.00	43.11
2174	O	HOH	1464	3.631	73.873	17.643	1.00	40.56
2175	O	HOH	1465	9.605	58.321	17.519	1.00	32.77
2176	O	HOH	1467	-17.032	68.963	3.058	1.00	55.63
2177	O	HOH	1468	3.971	90.021	-4.285	1.00	40.07
2178	O	HOH	1469	-2.156	92.850	4.908	1.00	28.34
2179	O	HOH	1470	-2.530	68.928	-5.556	1.00	45.41
2180	O	HOH	1471	10.068	55.463	-1.185	1.00	48.19
2181	O	HOH	1472	17.436	76.881	-8.410	1.00	39.74
2182	O	HOH	1473	21.118	70.834	-12.727	1.00	41.30
2183	O	HOH	1474	14.017	62.008	-12.371	1.00	37.31
2184	O	HOH	1475	3.766	64.917	-25.290	1.00	46.58
2185	O	HOH	1476	3.820	78.872	-19.171	1.00	37.77
2186	O	HOH	1477	14.054	67.327	14.991	1.00	47.15
2187	O	HOH	1478	16.414	66.571	8.050	1.00	33.09
2188	O	HOH	1479	13.492	90.782	8.574	1.00	35.84
2189	O	HOH	1480	-1.459	51.496	-2.038	1.00	35.40
2190	O	HOH	1481	-13.999	46.676	14.668	1.00	57.29
2191	O	HOH	1482	8.078	51.214	16.312	1.00	41.90
2192	O	HOH	1483	25.076	85.403	-10.681	1.00	42.14
2193	O	HOH	1484	14.827	42.032	13.274	1.00	38.01

2194	O	HOH	1485	5.345	91.858	-6.879	1.00	21.70
2195	O	HOH	1486	15.878	90.055	7.000	1.00	33.91
2196	O	HOH	1487	-3.296	75.542	21.607	1.00	45.24
2197	O	HOH	1488	19.842	69.177	-14.215	1.00	40.96
2198	O	HOH	1489	8.685	94.071	3.387	1.00	47.43
2199	O	HOH	1491	3.431	86.916	-5.716	1.00	36.57
2200	O	HOH	1492	17.033	64.110	8.961	1.00	67.29
2201	O	HOH	1493	17.906	88.820	6.427	1.00	30.45
2202	O	HOH	1494	-0.341	82.992	10.570	1.00	42.78
2203	O	HOH	1496	13.721	63.920	14.926	1.00	36.78
2204	O	HOH	1497	-6.719	44.417	9.331	1.00	58.77
2205	O	HOH	1498	17.887	88.899	9.606	1.00	46.91
2206	O	HOH	1499	-3.784	91.582	6.424	1.00	60.43
2207	O	HOH	1500	8.466	47.385	15.082	1.00	52.47
2208	O	HOH	1502	17.681	87.834	-23.073	1.00	55.73
2209	O	HOH	1504	22.284	88.287	-9.174	1.00	30.91
2210	O	HOH	1508	23.871	98.150	7.738	1.00	52.44
2211	O	HOH	1509	-14.428	56.398	6.034	1.00	45.65
2212	O	HOH	1510	11.490	49.327	-4.215	1.00	54.26
2213	O	HOH	1511	5.581	89.532	-6.136	1.00	42.45
2214	O	HOH	1512	-8.766	43.093	9.415	1.00	67.82
2215	O	HOH	1513	23.914	78.098	-9.885	1.00	43.63
2216	O	HOH	1514	9.169	49.305	1.558	1.00	53.26
2217	O	HOH	1515	0.623	84.681	-10.709	1.00	56.87
2218	O	HOH	1517	21.198	55.662	-7.695	1.00	41.44
2219	O	HOH	1518	13.949	49.147	-2.696	1.00	51.21
2220	O	HOH	1519	20.617	62.266	-5.635	1.00	44.95
2221	O	HOH	1521	17.504	69.377	-15.949	1.00	68.39
2222	O	HOH	1522	-1.374	78.824	-8.095	1.00	39.66
2223	O	HOH	1524	16.180	70.568	12.091	1.00	37.58
2224	O	HOH	1525	10.939	64.265	-14.059	1.00	47.37
2225	O	HOH	1526	-18.912	49.582	7.500	1.00	54.08
2226	O	HOH	1527	-14.931	60.138	2.590	1.00	39.16
2227	O	HOH	1528	3.061	74.939	-25.203	1.00	47.50
2228	O	HOH	1529	6.488	84.875	-15.940	1.00	43.46
2229	O	HOH	1531	15.738	73.049	11.537	1.00	46.88
2230	O	HOH	1532	5.127	83.103	-17.683	1.00	41.07
2231	O	HOH	1533	-7.085	44.550	6.909	1.00	58.24
2232	O	HOH	1535	16.269	80.299	12.472	1.00	58.48
2233	O	HOH	1536	17.373	84.555	-18.995	1.00	60.54
2234	O	HOH	1537	-2.998	87.771	6.384	1.00	38.46
2235	O	HOH	1538	-14.510	47.690	5.365	1.00	54.21
2236	O	HOH	1539	-0.950	56.797	22.634	1.00	54.75
2237	O	HOH	1540	8.668	80.241	12.792	1.00	58.85
2238	O	HOH	1541	-13.714	57.378	8.781	1.00	77.19
2239	O	HOH	1542	11.484	88.801	9.860	1.00	53.88
2240	O	HOH	1543	15.150	49.002	-5.155	1.00	57.17
2241	O	HOH	1544	-17.155	61.886	3.452	1.00	53.58
2242	O	HOH	1545	-4.290	70.619	-18.621	1.00	60.77
2243	O	HOH	1546	-7.645	81.391	14.039	1.00	67.28
2244	O	HOH	1547	-5.308	49.298	17.797	1.00	37.11
2245	O	HOH	1548	16.528	58.159	-0.465	1.00	36.55
2246	O	HOH	1549	-3.904	52.510	-3.165	1.00	52.07
2247	O	HOH	1550	12.314	94.588	-2.261	1.00	47.37
2248	O	HOH	1551	-19.562	69.135	10.528	1.00	60.08

2249	O	HOH	1552	2.866	75.832	8.852	1.00	54.95
2250	O	HOH	1553	-16.153	51.620	-2.755	1.00	40.11
2251	O	HOH	1554	-2.784	42.095	16.949	1.00	52.37
2252	O	HOH	1555	-6.866	59.219	22.388	1.00	65.60
2253	O	HOH	1556	23.298	100.259	6.422	1.00	55.17
2254	O	HOH	1557	15.280	88.036	-9.606	1.00	61.20
2255	O	HOH	1558	19.614	66.648	-15.163	1.00	67.89
2256	O	HOH	1559	-7.545	70.931	-4.844	1.00	44.74
2257	O	HOH	1560	-18.106	72.996	16.007	1.00	52.21
2258	O	HOH	1561	-6.264	65.696	-13.419	1.00	51.95
2259	O	HOH	1562	-2.205	92.669	8.048	1.00	51.77
2260	O	HOH	1563	0.621	56.635	20.299	1.00	43.44
2261	O	HOH	1564	1.924	71.413	-15.495	1.00	37.28
2262	O	HOH	1565	-6.379	46.446	5.260	1.00	40.01
2263	O	HOH	1566	18.514	87.780	-18.304	1.00	52.37
2264	O	HOH	1567	-9.892	82.981	13.612	1.00	52.74
2265	O	HOH	1568	25.016	73.814	-11.361	1.00	55.37
2266	O	HOH	1569	26.172	69.403	-6.808	1.00	54.99
2267	O	HOH	1570	-10.751	41.407	10.428	1.00	65.08
2268	O	HOH	1571	-1.525	44.394	18.040	1.00	39.22
2269	O	HOH	1572	-2.479	78.007	21.984	1.00	60.39
2270	O	HOH	1573	29.543	91.385	-7.198	1.00	56.92
2271	O	HOH	1574	-2.377	66.809	-20.717	1.00	48.89
2272	O	HOH	1575	-16.541	47.664	15.566	1.00	54.36
2273	O	HOH	1576	22.611	82.812	4.431	1.00	40.50
2274	O	HOH	1577	1.879	68.844	-15.655	1.00	43.48
2275	O	HOH	1579	-14.564	62.187	4.134	1.00	85.68
2276	O	HOH	1580	-0.922	68.511	-7.695	1.00	61.30
2277	O	HOH	1581	8.558	84.765	-12.374	1.00	69.19
2278	O	HOH	1582	6.658	78.557	14.489	1.00	49.03
2279	O	HOH	1583	30.956	87.513	-7.238	1.00	53.00
2280	O	HOH	1584	-17.862	50.491	-5.083	1.00	38.77
2281	O	HOH	1585	2.004	71.770	-26.632	1.00	58.07
2282	O	HOH	1586	12.168	88.002	-8.450	1.00	41.72
2283	O	HOH	1587	-0.542	67.095	-24.774	1.00	62.48
2284	O	HOH	1588	-1.879	62.615	23.873	1.00	56.77
2285	O	HOH	1589	1.344	80.543	10.432	1.00	55.28
2286	O	HOH	1590	8.362	60.245	18.594	1.00	57.98
2287	O	HOH	1591	-10.486	78.812	5.493	1.00	57.71
2288	O	HOH	1592	14.194	65.739	-20.629	1.00	64.34
2289	O	HOH	1593	18.091	65.351	-13.811	1.00	66.02
2290	O	HOH	1594	8.341	89.857	6.605	1.00	76.90
2291	C1	MON	1	7.260	75.391	0.223	0.40	52.85
2292	C2	MON	1	7.644	76.769	-0.053	0.40	52.34
2293	C3	MON	1	8.940	77.233	-0.542	0.40	51.77
2294	C4	MON	1	9.879	76.152	-0.742	0.40	53.59
2295	C5	MON	1	9.578	74.806	-0.497	0.40	55.34
2296	C6	MON	1	8.301	74.453	-0.023	0.40	55.00
2297	C7	MON	1	10.546	73.772	-0.742	0.40	57.15
2298	C8	MON	1	10.593	73.174	-1.905	0.40	59.86
2299	P9	MON	1	12.296	72.820	-2.589	0.40	61.27
2300	P10	MON	1	9.170	72.804	-2.984	0.40	60.74
2301	O11	MON	1	6.734	77.697	0.147	0.40	52.04
2302	C12	MON	1	5.878	74.909	0.740	0.40	51.45
2303	C13	MON	1	9.465	78.678	-0.870	0.40	51.45

2304	C14	MON	1	4.792	75.230	-0.335	0.40	51.26
2305	C15	MON	1	5.542	75.599	2.090	0.40	51.97
2306	C16	MON	1	5.764	73.344	0.977	0.40	49.94
2307	C17	MON	1	9.884	78.747	-2.388	0.40	50.76
2308	C18	MON	1	8.445	79.863	-0.657	0.40	49.85
2309	C19	MON	1	10.741	78.992	0.017	0.40	50.57
2310	O20	MON	1	12.425	71.309	-2.657	0.40	60.40
2311	O21	MON	1	13.385	73.256	-1.505	0.40	60.88
2312	C22	MON	1	14.690	73.453	-2.041	0.40	61.26
2313	C23	MON	1	15.662	73.866	-0.929	0.40	61.79
2314	O24	MON	1	12.591	73.280	-4.051	0.40	62.30
2315	C25	MON	1	12.208	74.590	-4.371	0.40	63.15
2316	C26	MON	1	12.571	74.879	-5.824	0.40	63.78
2317	O27	MON	1	7.838	72.942	-2.449	0.40	60.35
2318	O28	MON	1	9.304	73.724	-4.294	0.40	60.33
2319	C29	MON	1	8.099	74.504	-4.497	0.40	60.19
2320	C30	MON	1	8.203	75.367	-5.728	0.40	59.86
2321	O31	MON	1	9.396	71.343	-3.610	0.40	60.40
2322	C32	MON	1	9.336	70.349	-2.650	0.40	61.15
2323	C33	MON	1	9.565	69.036	-3.314	0.40	61.02
2324	C1	MON	4	12.842	73.050	-1.061	0.40	55.03
2325	C2	MON	4	11.639	72.852	-1.861	0.40	55.00
2326	C3	MON	4	10.384	73.606	-1.763	0.40	54.62
2327	C4	MON	4	10.420	74.643	-0.730	0.40	54.71
2328	C5	MON	4	11.551	74.898	0.085	0.40	54.65
2329	C6	MON	4	12.720	74.115	-0.086	0.40	55.01
2330	C7	MON	4	11.588	75.962	1.093	0.40	54.04
2331	C8	MON	4	11.306	77.220	0.796	0.40	54.50
2332	P9	MON	4	12.053	77.847	-0.850	0.40	53.52
2333	P10	MON	4	10.154	78.334	1.625	0.40	55.72
2334	O11	MON	4	11.687	71.897	-2.781	0.40	55.22
2335	C12	MON	4	14.157	72.233	-1.193	0.40	55.30
2336	C13	MON	4	9.029	73.487	-2.576	0.40	54.63
2337	C14	MON	4	14.744	72.410	-2.632	0.40	55.03
2338	C15	MON	4	13.871	70.729	-0.888	0.40	56.09
2339	C16	MON	4	15.323	72.668	-0.193	0.40	55.32
2340	C17	MON	4	8.742	74.858	-3.304	0.40	55.14
2341	C18	MON	4	8.973	72.376	-3.716	0.40	55.45
2342	C19	MON	4	7.840	73.192	-1.582	0.40	54.46
2343	O20	MON	4	11.010	78.626	-1.586	0.40	52.97
2344	O21	MON	4	12.374	76.619	-1.845	0.40	54.37
2345	C22	MON	4	13.541	76.839	-2.687	0.40	56.13
2346	C23	MON	4	13.810	75.656	-3.604	0.40	56.44
2347	O24	MON	4	13.248	78.838	-0.653	0.40	55.46
2348	C25	MON	4	14.184	78.428	0.368	0.40	55.77
2349	C26	MON	4	15.313	79.458	0.480	0.40	56.44
2350	O27	MON	4	9.716	78.033	2.935	0.40	56.39
2351	O28	MON	4	10.783	79.845	1.379	0.40	55.48
2352	C29	MON	4	10.133	80.538	0.242	0.40	56.37
2353	C30	MON	4	10.735	81.909	0.029	0.40	57.08
2354	O31	MON	4	8.770	78.294	0.725	0.40	55.95
2355	C32	MON	4	8.067	77.052	0.863	0.40	55.63
2356	C33	MON	4	6.756	77.095	0.032	0.40	56.56
2357	C1	MON	5	12.648	74.791	0.798	0.20	47.05
2358	C2	MON	5	13.126	76.119	0.932	0.20	47.32

2359	C3	MON	5	12.656	77.303	0.159	0.20	47.22
2360	C4	MON	5	11.616	76.953	-0.780	0.20	47.19
2361	C5	MON	5	11.126	75.646	-0.931	0.20	47.61
2362	C6	MON	5	11.639	74.637	-0.165	0.20	47.35
2363	C7	MON	5	10.111	75.325	-1.848	0.20	47.58
2364	C8	MON	5	8.874	74.929	-1.501	0.20	48.33
2365	P9	MON	5	7.588	75.119	-2.808	0.20	47.86
2366	P10	MON	5	8.265	74.266	0.100	0.20	50.66
2367	O11	MON	5	14.069	76.314	1.815	0.20	47.52
2368	C12	MON	5	13.108	73.595	1.565	0.20	46.60
2369	C13	MON	5	13.080	78.852	0.183	0.20	47.46
2370	C14	MON	5	12.817	73.796	3.086	0.20	45.89
2371	C15	MON	5	14.636	73.345	1.297	0.20	46.19
2372	C16	MON	5	12.362	72.280	1.154	0.20	46.10
2373	C17	MON	5	11.819	79.734	0.589	0.20	47.01
2374	C18	MON	5	14.218	79.266	1.196	0.20	47.12
2375	C19	MON	5	13.555	79.299	-1.261	0.20	47.39
2376	O20	MON	5	6.243	75.028	-2.117	0.20	47.51
2377	O21	MON	5	7.612	73.832	-3.795	0.20	47.61
2378	C22	MON	5	8.896	73.594	-4.380	0.20	47.53
2379	C23	MON	5	8.868	72.365	-5.290	0.20	47.07
2380	O24	MON	5	7.601	76.547	-3.505	0.20	48.11
2381	C25	MON	5	8.083	77.621	-2.598	0.20	48.54
2382	C26	MON	5	8.106	78.979	-3.250	0.20	48.93
2383	O27	MON	5	9.172	74.062	1.195	0.20	50.80
2384	O28	MON	5	6.985	75.197	0.526	0.20	50.30
2385	C29	MON	5	7.371	76.386	1.227	0.20	50.25
2386	C30	MON	5	6.133	77.218	1.594	0.20	49.96
2387	O31	MON	5	7.569	72.853	-0.242	0.20	49.96
2388	C32	MON	5	8.536	71.922	-0.647	0.20	49.66
2389	C33	MON	5	7.882	70.608	-0.958	0.20	49.90
2390	O	HOH	1595	7.323	78.531	4.526	1.00	69.10
2391	O	HOH	1596	17.213	76.326	0.429	1.00	57.74
2392	O	HOH	1597	-12.947	67.694	7.441	1.00	77.03
2393	O	HOH	1598	16.163	67.413	-22.431	1.00	40.69
2394	O	HOH	1599	18.369	92.512	7.265	1.00	52.48
2395	O	HOH	1600	29.442	64.461	-1.726	1.00	66.79
2396	O	HOH	1601	19.427	85.921	-22.662	1.00	60.16
2397	O	HOH	1602	5.344	90.815	7.154	1.00	54.96
2398	O	HOH	1603	-14.216	50.571	5.561	1.00	54.96
2399	O	HOH	1604	5.533	45.964	0.404	1.00	62.55
2400	O	HOH	1605	-1.394	63.145	20.705	1.00	40.08
2401	O	HOH	1606	-2.578	54.566	22.874	1.00	57.40
2402	O	HOH	1607	3.600	69.196	22.807	1.00	54.51
2403	O	HOH	1608	6.139	65.007	-18.611	1.00	54.86
2404	O	HOH	1609	4.202	75.224	-27.568	1.00	58.04
2405	O	HOH	1610	-5.421	61.703	24.061	1.00	57.88
2406	O	HOH	1611	-11.943	45.372	11.041	1.00	62.72

TABLE 5

ATOMIC COORDINATE DATA FOR HUMAN VITAMIN D RECEPTOR EMPLOYED IN THE MOLECULAR REPLACEMENT SOLUTION OF HUMAN PXR LIGAND BINDING DOMAIN CRYSTALS

ATOM	ATOM TYPE	RESIDUE	#	X	Y	Z	OCC	B	ATOM
1	N	LEU	A	120	34.417	18.787	67.312	1.00	50.31
2	CA	LEU	A	120	34.298	17.304	67.212	1.00	49.96
3	C	LEU	A	120	33.672	16.891	65.886	1.00	49.44
4	O	LEU	A	120	32.815	17.592	65.344	1.00	49.49
5	CB	LEU	A	120	33.447	16.756	68.363	1.00	50.64
9	N	ARG	A	121	34.106	15.745	65.375	1.00	48.14
10	CA	ARG	A	121	33.599	15.221	64.117	1.00	47.01
11	C	ARG	A	121	33.113	13.790	64.314	1.00	45.50
12	O	ARG	A	121	33.775	12.836	63.905	1.00	45.36
13	CB	ARG	A	121	34.700	15.264	63.052	1.00	48.45
14	CG	ARG	A	121	35.233	16.664	62.790	1.00	49.89
15	CD	ARG	A	121	36.430	16.655	61.852	1.00	52.32
16	NE	ARG	A	121	36.100	16.133	60.529	1.00	53.49
17	CZ	ARG	A	121	36.947	16.112	59.504	1.00	54.08
18	NH1	ARG	A	121	38.178	16.586	59.648	1.00	54.50
19	NH2	ARG	A	121	36.563	15.620	58.334	1.00	54.12
20	N	PRO	A	122	31.946	13.622	64.955	1.00	43.87
21	CA	PRO	A	122	31.403	12.282	65.187	1.00	42.99
22	C	PRO	A	122	31.173	11.529	63.881	1.00	42.25
23	O	PRO	A	122	30.823	12.125	62.862	1.00	42.01
24	CB	PRO	A	122	30.105	12.561	65.944	1.00	42.59
27	N	LYS	A	123	31.379	10.218	63.920	1.00	41.53
28	CA	LYS	A	123	31.205	9.378	62.744	1.00	41.30
29	C	LYS	A	123	29.732	9.158	62.431	1.00	40.35
30	O	LYS	A	123	28.877	9.250	63.313	1.00	39.21
31	CB	LYS	A	123	31.885	8.024	62.965	1.00	42.56
36	N	LEU	A	124	29.439	8.879	61.165	1.00	39.48
37	CA	LEU	A	124	28.071	8.622	60.744	1.00	38.64
38	C	LEU	A	124	27.606	7.325	61.384	1.00	38.41
39	O	LEU	A	124	28.293	6.308	61.304	1.00	39.12
40	CB	LEU	A	124	27.996	8.491	59.220	1.00	37.76
41	CG	LEU	A	124	28.162	9.776	58.406	1.00	37.83
42	CD1	LEU	A	124	28.401	9.438	56.941	1.00	37.98
43	CD2	LEU	A	124	26.922	10.633	58.564	1.00	36.97
44	N	SER	A	125	26.448	7.362	62.029	1.00	38.47
45	CA	SER	A	125	25.905	6.168	62.661	1.00	39.40
46	C	SER	A	125	25.496	5.197	61.561	1.00	40.52
47	O	SER	A	125	25.386	5.581	60.395	1.00	39.53
48	CB	SER	A	125	24.679	6.523	63.495	1.00	39.88
50	N	GLU	A	126	25.271	3.940	61.923	1.00	41.33
51	CA	GLU	A	126	24.865	2.956	60.930	1.00	42.41

52	C	GLU	A	126	23.535	3.385	60.314	1.00	41.49
53	O	GLU	A	126	23.313	3.207	59.115	1.00	41.40
54	CB	GLU	A	126	24.727	1.573	61.573	1.00	45.02
55	CG	GLU	A	126	24.325	0.463	60.605	1.00	48.95
56	CD	GLU	A	126	25.202	0.414	59.361	1.00	51.93
57	OE1	GLU	A	126	24.878	1.105	58.366	1.00	53.34
58	OE2	GLU	A	126	26.222	-0.308	59.379	1.00	53.64
59	N	GLU	A	127	22.659	3.960	61.133	1.00	40.27
60	CA	GLU	A	127	21.358	4.412	60.650	1.00	39.52
61	C	GLU	A	127	21.512	5.550	59.647	1.00	37.38
62	O	GLU	A	127	20.814	5.594	58.630	1.00	36.24
63	CB	GLU	A	127	20.481	4.891	61.807	1.00	41.53
64	CG	GLU	A	127	19.091	5.320	61.363	1.00	45.78
65	CD	GLU	A	127	18.236	5.832	62.504	1.00	47.87
66	OE1	GLU	A	127	18.572	6.890	63.075	1.00	49.93
67	OE2	GLU	A	127	17.227	5.173	62.832	1.00	50.45
68	N	GLN	A	128	22.420	6.473	59.939	1.00	34.92
69	CA	GLN	A	128	22.654	7.603	59.052	1.00	33.94
70	C	GLN	A	128	23.239	7.134	57.721	1.00	34.19
71	O	GLN	A	128	22.905	7.671	56.665	1.00	32.45
72	CB	GLN	A	128	23.573	8.622	59.735	1.00	33.20
73	CG	GLN	A	128	22.861	9.410	60.835	1.00	32.00
74	CD	GLN	A	128	23.785	10.317	61.629	1.00	32.20
75	OE1	GLN	A	128	23.346	11.326	62.192	1.00	33.66
76	NE2	GLN	A	128	25.061	9.960	61.691	1.00	30.80
77	N	GLN	A	129	24.101	6.124	57.768	1.00	33.75
78	CA	GLN	A	129	24.692	5.591	56.545	1.00	35.00
79	C	GLN	A	129	23.588	4.965	55.702	1.00	34.31
80	O	GLN	A	129	23.562	5.111	54.479	1.00	33.78
81	CB	GLN	A	129	25.747	4.531	56.874	1.00	37.89
86	N	ARG	A	130	22.674	4.270	56.370	1.00	33.44
87	CA	ARG	A	130	21.556	3.614	55.703	1.00	34.05
88	C	ARG	A	130	20.653	4.638	55.018	1.00	32.98
89	O	ARG	A	130	20.226	4.436	53.881	1.00	31.44
90	CB	ARG	A	130	20.759	2.794	56.723	1.00	37.04
97	N	ILE	A	131	20.367	5.735	55.712	1.00	31.16
98	CA	ILE	A	131	19.519	6.790	55.158	1.00	30.41
99	C	ILE	A	131	20.120	7.343	53.865	1.00	29.21
100	O	ILE	A	131	19.414	7.528	52.872	1.00	27.86
101	CB	ILE	A	131	19.334	7.945	56.177	1.00	31.61
105	N	ILE	A	132	21.424	7.601	53.876	1.00	28.81
106	CA	ILE	A	132	22.094	8.124	52.691	1.00	29.13
107	C	ILE	A	132	22.029	7.115	51.544	1.00	29.37
108	O	ILE	A	132	21.786	7.486	50.394	1.00	28.72
109	CB	ILE	A	132	23.570	8.468	52.994	1.00	29.90
110	CG1	ILE	A	132	23.628	9.625	53.995	1.00	30.31
111	CG2	ILE	A	132	24.306	8.838	51.708	1.00	30.32
112	CD1	ILE	A	132	25.027	9.997	54.432	1.00	31.33
113	N	ALA	A	133	22.239	5.841	51.862	1.00	28.31
114	CA	ALA	A	133	22.203	4.785	50.851	1.00	27.51
115	C	ALA	A	133	20.820	4.680	50.213	1.00	26.94
116	O	ALA	A	133	20.694	4.542	48.993	1.00	26.91
117	CB	ALA	A	133	22.587	3.454	51.479	1.00	27.94
118	N	ILE	A	134	19.786	4.739	51.044	1.00	26.00
119	CA	ILE	A	134	18.413	4.659	50.564	1.00	25.19

120	C	ILE	A	134	18.090	5.832	49.643	1.00	24.84
121	O	ILE	A	134	17.490	5.651	48.585	1.00	23.10
122	CB	ILE	A	134	17.416	4.660	51.742	1.00	26.47
126	N	LEU	A	135	18.494	7.030	50.047	1.00	23.54
127	CA	LEU	A	135	18.228	8.220	49.242	1.00	23.28
128	C	LEU	A	135	18.987	8.217	47.914	1.00	22.05
129	O	LEU	A	135	18.454	8.656	46.894	1.00	21.44
130	CB	LEU	A	135	18.559	9.480	50.045	1.00	23.21
131	CG	LEU	A	135	17.644	9.754	51.246	1.00	24.57
132	CD1	LEU	A	135	18.057	11.076	51.900	1.00	26.44
133	CD2	LEU	A	135	16.185	9.820	50.789	1.00	25.56
134	N	LEU	A	136	20.223	7.725	47.913	1.00	22.40
135	CA	LEU	A	136	20.991	7.675	46.669	1.00	23.29
136	C	LEU	A	136	20.302	6.721	45.705	1.00	23.50
137	O	LEU	A	136	20.191	6.996	44.512	1.00	23.31
138	CB	LEU	A	136	22.424	7.194	46.920	1.00	24.60
142	N	ASP	A	137	19.845	5.591	46.232	1.00	23.87
143	CA	ASP	A	137	19.156	4.589	45.427	1.00	23.95
144	C	ASP	A	137	17.844	5.152	44.870	1.00	23.67
145	O	ASP	A	137	17.513	4.943	43.697	1.00	22.79
146	CB	ASP	A	137	18.886	3.348	46.282	1.00	26.93
147	CG	ASP	A	137	18.158	2.266	45.524	1.00	31.10
148	OD1	ASP	A	137	17.010	1.947	45.900	1.00	34.78
149	OD2	ASP	A	137	18.730	1.734	44.552	1.00	34.13
150	N	ALA	A	138	17.105	5.867	45.714	1.00	22.31
151	CA	ALA	A	138	15.836	6.472	45.312	1.00	22.31
152	C	ALA	A	138	16.063	7.435	44.157	1.00	21.39
153	O	ALA	A	138	15.310	7.445	43.183	1.00	20.83
154	CB	ALA	A	138	15.213	7.219	46.487	1.00	23.04
155	N	HIS	A	139	17.107	8.249	44.263	1.00	21.06
156	CA	HIS	A	139	17.408	9.202	43.208	1.00	21.28
157	C	HIS	A	139	17.814	8.511	41.905	1.00	21.64
158	O	HIS	A	139	17.385	8.913	40.824	1.00	21.17
159	CB	HIS	A	139	18.528	10.152	43.631	1.00	21.21
165	N	HIS	A	140	18.650	7.479	42.005	1.00	21.50
166	CA	HIS	A	140	19.099	6.760	40.819	1.00	22.20
167	C	HIS	A	140	17.947	6.088	40.082	1.00	21.95
168	O	HIS	A	140	17.997	5.911	38.861	1.00	21.87
169	CB	HIS	A	140	20.153	5.710	41.193	1.00	23.76
175	N	LYS	A	141	16.908	5.719	40.821	1.00	20.41
176	CA	LYS	A	141	15.745	5.071	40.225	1.00	21.89
177	C	LYS	A	141	14.746	6.078	39.665	1.00	21.31
178	O	LYS	A	141	13.916	5.730	38.832	1.00	22.47
179	CB	LYS	A	141	15.031	4.203	41.265	1.00	23.28
180	CG	LYS	A	141	15.804	2.960	41.668	1.00	26.83
181	CD	LYS	A	141	15.080	2.209	42.771	1.00	30.63
182	CE	LYS	A	141	15.781	0.902	43.093	1.00	33.64
183	NZ	LYS	A	141	15.122	0.206	44.231	1.00	36.58
184	N	THR	A	142	14.840	7.325	40.107	1.00	20.65
185	CA	THR	A	142	13.893	8.348	39.664	1.00	20.68
186	C	THR	A	142	14.440	9.502	38.833	1.00	20.45
187	O	THR	A	142	13.682	10.375	38.420	1.00	20.32
188	CB	THR	A	142	13.142	8.935	40.865	1.00	20.48
189	OG1	THR	A	142	14.081	9.474	41.805	1.00	18.91
190	CG2	THR	A	142	12.326	7.850	41.546	1.00	19.94

191	N	TYR	A	143	15.747	9.520	38.595	1.00	20.03
192	CA	TYR	A	143	16.342	10.566	37.768	1.00	20.44
193	C	TYR	A	143	17.207	9.895	36.706	1.00	20.75
194	O	TYR	A	143	18.248	9.323	37.013	1.00	21.56
195	CB	TYR	A	143	17.198	11.529	38.610	1.00	20.88
203	N	ASP	A	144	16.750	9.959	35.461	1.00	20.48
204	CA	ASP	A	144	17.449	9.365	34.326	1.00	21.36
205	C	ASP	A	144	18.428	10.387	33.751	1.00	22.06
206	O	ASP	A	144	18.016	11.348	33.102	1.00	21.75
207	CB	ASP	A	144	16.412	8.955	33.274	1.00	21.65
208	CG	ASP	A	144	17.032	8.481	31.976	1.00	22.22
209	OD1	ASP	A	144	18.261	8.286	31.921	1.00	22.12
210	OD2	ASP	A	144	16.266	8.294	31.007	1.00	23.20
211	N	PRO	A	145	19.741	10.183	33.976	1.00	21.93
212	CA	PRO	A	145	20.779	11.094	33.483	1.00	23.05
213	C	PRO	A	145	20.968	11.106	31.968	1.00	22.50
214	O	PRO	A	145	21.754	11.906	31.451	1.00	23.61
215	CB	PRO	A	145	22.026	10.620	34.225	1.00	23.45
218	N	THR	A	146	20.265	10.224	31.256	1.00	22.03
219	CA	THR	A	146	20.364	10.192	29.796	1.00	21.95
220	C	THR	A	146	19.174	10.907	29.155	1.00	22.52
221	O	THR	A	146	19.181	11.177	27.953	1.00	22.17
222	CB	THR	A	146	20.433	8.750	29.233	1.00	21.96
223	OG1	THR	A	146	19.167	8.099	29.395	1.00	21.08
224	CG2	THR	A	146	21.509	7.949	29.956	1.00	23.14
225	N	TYR	A	147	18.158	11.210	29.963	1.00	22.04
226	CA	TYR	A	147	16.963	11.912	29.489	1.00	22.53
227	C	TYR	A	147	16.313	11.191	28.309	1.00	23.10
228	O	TYR	A	147	15.789	11.821	27.393	1.00	23.05
229	CB	TYR	A	147	17.335	13.350	29.093	1.00	23.34
237	N	SER	A	148	16.326	9.862	28.355	1.00	23.29
238	CA	SER	A	148	15.781	9.046	27.278	1.00	23.65
239	C	SER	A	148	14.263	9.078	27.073	1.00	24.65
240	O	SER	A	148	13.783	8.650	26.024	1.00	24.62
241	CB	SER	A	148	16.243	7.593	27.450	1.00	26.66
242	OG	SER	A	148	15.684	7.006	28.614	1.00	29.82
243	N	ASP	A	149	13.505	9.576	28.048	1.00	22.99
244	CA	ASP	A	149	12.045	9.632	27.905	1.00	23.85
245	C	ASP	A	149	11.534	10.925	27.272	1.00	24.00
246	O	ASP	A	149	10.371	11.008	26.879	1.00	24.41
247	CB	ASP	A	149	11.349	9.488	29.263	1.00	24.47
251	N	PHE	A	150	12.396	11.927	27.171	1.00	24.31
252	CA	PHE	A	150	11.995	13.231	26.646	1.00	25.09
253	C	PHE	A	150	11.363	13.263	25.252	1.00	25.91
254	O	PHE	A	150	10.565	14.155	24.949	1.00	25.61
255	CB	PHE	A	150	13.188	14.187	26.715	1.00	24.68
256	CG	PHE	A	150	13.546	14.611	28.121	1.00	25.17
257	CD1	PHE	A	150	13.422	13.726	29.187	1.00	25.54
258	CD2	PHE	A	150	14.028	15.891	28.374	1.00	26.43
259	CE1	PHE	A	150	13.773	14.104	30.484	1.00	25.74
260	CE2	PHE	A	150	14.384	16.278	29.667	1.00	25.55
261	CZ	PHE	A	150	14.256	15.386	30.721	1.00	24.63
262	N	CYS	A	151	11.694	12.298	24.404	1.00	27.60
263	CA	CYS	A	151	11.116	12.286	23.063	1.00	28.74
264	C	CYS	A	151	9.640	11.891	23.094	1.00	28.90

265	O	CYS	A	151	8.951	11.958	22.075	1.00	28.40
266	CB	CYS	A	151	11.894	11.332	22.154	1.00	31.34
268	N	GLN	A	152	9.152	11.482	24.262	1.00	27.55
269	CA	GLN	A	152	7.753	11.093	24.393	1.00	27.93
270	C	GLN	A	152	6.858	12.285	24.711	1.00	27.73
271	O	GLN	A	152	5.633	12.202	24.590	1.00	28.51
272	CB	GLN	A	152	7.602	10.021	25.473	1.00	29.61
277	N	PHE	A	153	7.469	13.395	25.115	1.00	25.45
278	CA	PHE	A	153	6.705	14.597	25.439	1.00	25.30
279	C	PHE	A	153	6.261	15.273	24.151	1.00	25.61
280	O	PHE	A	153	6.799	14.998	23.071	1.00	24.69
281	CB	PHE	A	153	7.564	15.608	26.215	1.00	23.94
282	CG	PHE	A	153	8.187	15.060	27.469	1.00	23.45
283	CD1	PHE	A	153	9.332	15.654	27.990	1.00	22.75
284	CD2	PHE	A	153	7.654	13.949	28.116	1.00	23.40
285	CE1	PHE	A	153	9.948	15.146	29.133	1.00	23.18
286	CE2	PHE	A	153	8.261	13.434	29.263	1.00	22.50
287	CZ	PHE	A	153	9.414	14.037	29.769	1.00	22.91
288	N	ARG	A	154	5.276	16.158	24.260	1.00	25.51
289	CA	ARG	A	154	4.842	16.902	23.092	1.00	26.08
290	C	ARG	A	154	6.094	17.673	22.689	1.00	27.20
291	O	ARG	A	154	6.824	18.184	23.542	1.00	26.99
292	CB	ARG	A	154	3.681	17.830	23.449	1.00	26.73
293	CG	ARG	A	154	2.351	17.087	23.522	1.00	27.85
294	CD	ARG	A	154	1.232	17.964	24.066	1.00	27.71
295	NE	ARG	A	154	1.347	18.138	25.509	1.00	27.14
296	CZ	ARG	A	154	0.497	18.839	26.248	1.00	28.47
297	NH1	ARG	A	154	-0.538	19.444	25.677	1.00	29.16
298	NH2	ARG	A	154	0.673	18.919	27.560	1.00	27.66
299	N	PRO	A	155	6.368	17.757	21.384	1.00	27.28
300	CA	PRO	A	155	7.554	18.454	20.892	1.00	28.12
301	C	PRO	A	155	7.709	19.929	21.217	1.00	28.41
302	O	PRO	A	155	6.733	20.676	21.291	1.00	27.77
303	CB	PRO	A	155	7.491	18.206	19.388	1.00	28.83
306	N	PRO	A	156	8.956	20.361	21.437	1.00	28.25
307	CA	PRO	A	156	9.202	21.768	21.739	1.00	29.56
308	C	PRO	A	156	9.054	22.532	20.425	1.00	30.08
309	O	PRO	A	156	9.483	22.054	19.371	1.00	30.96
310	CB	PRO	A	156	10.640	21.763	22.250	1.00	29.92
311	CG	PRO	A	156	11.262	20.646	21.476	1.00	30.45
312	CD	PRO	A	156	10.198	19.573	21.538	1.00	29.15
313	N	VAL	A	157	8.417	23.693	20.489	1.00	30.75
314	CA	VAL	A	157	8.220	24.538	19.319	1.00	31.52
315	C	VAL	A	157	8.764	25.907	19.692	1.00	32.33
316	O	VAL	A	157	8.361	26.482	20.698	1.00	33.09
317	CB	VAL	A	157	6.727	24.663	18.962	1.00	31.97
318	CG1	VAL	A	157	6.544	25.654	17.825	1.00	32.48
319	CG2	VAL	A	157	6.177	23.302	18.573	1.00	32.24
320	N	ARG	A	158	9.681	26.425	18.885	1.00	33.83
321	CA	ARG	A	158	10.289	27.716	19.173	1.00	36.19
322	C	ARG	A	158	10.020	28.766	18.096	1.00	38.44
323	O	ARG	A	158	10.763	28.881	17.123	1.00	39.20
324	CB	ARG	A	158	11.794	27.523	19.367	1.00	35.86
325	CG	ARG	A	158	12.131	26.585	20.524	1.00	34.74
326	CD	ARG	A	158	13.606	26.231	20.561	1.00	35.06

327	NE	ARG	A	158	13.991	25.641	21.841	1.00	32.63
328	CZ	ARG	A	158	14.006	24.339	22.113	1.00	31.82
329	NH1	ARG	A	158	13.658	23.450	21.192	1.00	32.10
330	NH2	ARG	A	158	14.370	23.926	23.319	1.00	29.69
331	N	VAL	A	159	8.949	29.531	18.284	1.00	40.67
332	CA	VAL	A	159	8.568	30.574	17.338	1.00	42.44
333	C	VAL	A	159	9.511	31.767	17.432	1.00	43.24
334	O	VAL	A	159	10.170	31.968	18.451	1.00	42.85
335	CB	VAL	A	159	7.135	31.066	17.607	1.00	42.85
338	N	ASN	A	160	9.576	32.557	16.365	1.00	44.06
339	CA	ASN	A	160	10.440	33.730	16.357	1.00	44.92
340	C	ASN	A	160	9.876	34.768	17.320	1.00	45.24
341	O	ASN	A	160	8.728	35.198	17.185	1.00	45.27
342	CB	ASN	A	160	10.530	34.326	14.949	1.00	46.00
346	N	ASP	A	161	10.688	35.156	18.298	1.00	45.02
347	CA	ASP	A	161	10.282	36.142	19.289	1.00	44.79
348	C	ASP	A	161	11.515	36.834	19.862	1.00	44.74
349	O	ASP	A	161	11.679	36.939	21.077	1.00	44.64
350	CB	ASP	A	161	9.483	35.463	20.406	1.00	44.26
354	N	GLY	A	162	12.383	37.304	18.972	1.00	44.73
355	CA	GLY	A	162	13.592	37.977	19.409	1.00	44.74
356	C	GLY	A	162	13.292	39.196	20.261	1.00	44.56
357	O	GLY	A	162	14.135	39.638	21.042	1.00	45.10
358	N	GLY	A	163	12.086	39.736	20.116	1.00	44.30
359	CA	GLY	A	163	11.706	40.911	20.879	1.00	43.74
360	C	GLY	A	163	11.206	40.618	22.282	1.00	43.23
361	O	GLY	A	163	11.066	41.533	23.096	1.00	43.53
362	N	GLY	A	164	10.946	39.346	22.572	1.00	42.43
363	CA	GLY	A	164	10.450	38.980	23.889	1.00	40.70
364	C	GLY	A	164	9.094	39.616	24.130	1.00	39.47
365	O	GLY	A	164	8.812	40.125	25.222	1.00	40.10
366	N	SER	A	216	8.256	39.587	23.099	1.00	36.82
367	CA	SER	A	216	6.918	40.165	23.162	1.00	35.37
368	C	SER	A	216	5.965	39.359	24.032	1.00	34.15
369	O	SER	A	216	5.653	38.213	23.721	1.00	32.50
370	CB	SER	A	216	6.329	40.277	21.755	1.00	35.39
372	N	VAL	A	217	5.495	39.969	25.116	1.00	33.39
373	CA	VAL	A	217	4.563	39.301	26.013	1.00	33.22
374	C	VAL	A	217	3.299	38.922	25.251	1.00	32.19
375	O	VAL	A	217	2.783	37.816	25.399	1.00	31.92
376	CB	VAL	A	217	4.161	40.208	27.195	1.00	33.21
377	CG1	VAL	A	217	3.203	39.462	28.119	1.00	35.52
378	CG2	VAL	A	217	5.396	40.644	27.960	1.00	35.70
379	N	THR	A	218	2.809	39.846	24.428	1.00	31.30
380	CA	THR	A	218	1.597	39.609	23.653	1.00	30.58
381	C	THR	A	218	1.736	38.398	22.741	1.00	30.30
382	O	THR	A	218	0.852	37.544	22.695	1.00	30.29
383	CB	THR	A	218	1.235	40.843	22.802	1.00	30.65
386	N	LEU	A	219	2.849	38.325	22.018	1.00	29.44
387	CA	LEU	A	219	3.095	37.206	21.117	1.00	29.87
388	C	LEU	A	219	3.260	35.905	21.894	1.00	29.21
389	O	LEU	A	219	2.710	34.869	21.516	1.00	29.73
390	CB	LEU	A	219	4.355	37.462	20.286	1.00	31.48
391	CG	LEU	A	219	4.778	36.321	19.352	1.00	33.59
392	CD1	LEU	A	219	3.700	36.083	18.301	1.00	34.93

393	CD2	LEU	A	219	6.100	36.676	18.690	1.00	35.57
394	N	GLU	A	220	4.018	35.963	22.982	1.00	28.82
395	CA	GLU	A	220	4.258	34.781	23.801	1.00	29.09
396	C	GLU	A	220	2.958	34.194	24.342	1.00	29.07
397	O	GLU	A	220	2.757	32.983	24.297	1.00	27.80
398	CB	GLU	A	220	5.213	35.131	24.946	1.00	31.33
403	N	LEU	A	221	2.073	35.052	24.841	1.00	28.79
404	CA	LEU	A	221	0.799	34.592	25.383	1.00	29.82
405	C	LEU	A	221	-0.143	34.089	24.293	1.00	29.77
406	O	LEU	A	221	-0.923	33.165	24.516	1.00	30.04
407	CB	LEU	A	221	0.125	35.714	26.181	1.00	30.05
408	CG	LEU	A	221	0.743	36.046	27.544	1.00	31.65
409	CD1	LEU	A	221	0.065	37.278	28.138	1.00	32.22
410	CD2	LEU	A	221	0.588	34.850	28.482	1.00	31.89
411	N	SER	A	222	-0.066	34.687	23.108	1.00	31.28
412	CA	SER	A	222	-0.931	34.272	22.011	1.00	32.25
413	C	SER	A	222	-0.536	32.905	21.460	1.00	32.84
414	O	SER	A	222	-1.380	32.170	20.947	1.00	33.76
415	CB	SER	A	222	-0.895	35.304	20.877	1.00	34.81
417	N	GLN	A	223	0.742	32.558	21.584	1.00	31.84
418	CA	GLN	A	223	1.234	31.288	21.063	1.00	31.75
419	C	GLN	A	223	1.596	30.215	22.089	1.00	30.53
420	O	GLN	A	223	1.306	29.039	21.869	1.00	30.69
421	CB	GLN	A	223	2.434	31.550	20.151	1.00	34.71
426	N	LEU	A	224	2.226	30.610	23.195	1.00	28.64
427	CA	LEU	A	224	2.632	29.654	24.232	1.00	27.07
428	C	LEU	A	224	3.209	28.401	23.569	1.00	26.40
429	O	LEU	A	224	2.898	27.274	23.962	1.00	25.81
430	CB	LEU	A	224	1.424	29.276	25.102	1.00	27.70
434	N	SER	A	225	4.071	28.614	22.577	1.00	25.74
435	CA	SER	A	225	4.667	27.531	21.798	1.00	25.83
436	C	SER	A	225	5.454	26.473	22.563	1.00	25.18
437	O	SER	A	225	5.446	25.302	22.182	1.00	25.89
438	CB	SER	A	225	5.557	28.110	20.696	1.00	26.31
439	OG	SER	A	225	6.710	28.731	21.233	1.00	29.36
440	N	MET	A	226	6.132	26.880	23.630	1.00	24.58
441	CA	MET	A	226	6.931	25.948	24.424	1.00	24.51
442	C	MET	A	226	6.193	25.387	25.631	1.00	24.00
443	O	MET	A	226	6.725	24.529	26.336	1.00	24.19
444	CB	MET	A	226	8.219	26.629	24.905	1.00	24.70
448	N	LEU	A	227	4.969	25.850	25.872	1.00	23.05
449	CA	LEU	A	227	4.225	25.377	27.030	1.00	23.85
450	C	LEU	A	227	3.882	23.887	27.032	1.00	23.65
451	O	LEU	A	227	4.062	23.218	28.052	1.00	24.44
452	CB	LEU	A	227	2.949	26.212	27.237	1.00	24.02
453	CG	LEU	A	227	2.139	25.868	28.494	1.00	24.67
454	CD1	LEU	A	227	3.019	25.994	29.730	1.00	25.75
455	CD2	LEU	A	227	0.936	26.798	28.612	1.00	25.81
456	N	PRO	A	228	3.395	23.336	25.901	1.00	24.00
457	CA	PRO	A	228	3.073	21.904	25.931	1.00	23.78
458	C	PRO	A	228	4.261	21.024	26.330	1.00	23.69
459	O	PRO	A	228	4.123	20.109	27.155	1.00	23.20
460	CB	PRO	A	228	2.602	21.626	24.504	1.00	24.23
461	CG	PRO	A	228	1.957	22.939	24.110	1.00	24.58
462	CD	PRO	A	228	2.962	23.948	24.629	1.00	23.63

463	N	HIS	A	229	5.421	21.305	25.747	1.00	22.38
464	CA	HIS	A	229	6.626	20.532	26.037	1.00	22.16
465	C	HIS	A	229	7.089	20.679	27.490	1.00	21.32
466	O	HIS	A	229	7.409	19.687	28.151	1.00	20.38
467	CB	HIS	A	229	7.765	20.951	25.103	1.00	22.65
468	CG	HIS	A	229	9.037	20.196	25.337	1.00	23.54
469	ND1	HIS	A	229	9.235	18.910	24.883	1.00	24.88
470	CD2	HIS	A	229	10.160	20.535	26.012	1.00	23.99
471	CE1	HIS	A	229	10.427	18.488	25.270	1.00	25.42
472	NE2	HIS	A	229	11.009	19.455	25.957	1.00	23.32
473	N	LEU	A	230	7.139	21.913	27.985	1.00	20.90
474	CA	LEU	A	230	7.578	22.139	29.355	1.00	21.22
475	C	LEU	A	230	6.563	21.623	30.361	1.00	21.08
476	O	LEU	A	230	6.938	21.164	31.435	1.00	19.50
477	CB	LEU	A	230	7.858	23.625	29.602	1.00	21.98
481	N	ALA	A	231	5.279	21.703	30.022	1.00	20.40
482	CA	ALA	A	231	4.243	21.197	30.917	1.00	21.50
483	C	ALA	A	231	4.421	19.685	31.040	1.00	21.12
484	O	ALA	A	231	4.303	19.124	32.129	1.00	21.78
485	CB	ALA	A	231	2.859	21.522	30.361	1.00	22.95
486	N	ASP	A	232	4.707	19.028	29.919	1.00	20.87
487	CA	ASP	A	232	4.910	17.582	29.916	1.00	21.48
488	C	ASP	A	232	6.168	17.228	30.711	1.00	20.43
489	O	ASP	A	232	6.167	16.259	31.463	1.00	21.59
490	CB	ASP	A	232	5.022	17.056	28.482	1.00	21.87
491	CG	ASP	A	232	3.664	16.893	27.807	1.00	25.14
492	OD1	ASP	A	232	3.639	16.665	26.582	1.00	26.65
493	OD2	ASP	A	232	2.623	16.982	28.497	1.00	25.49
494	N	LEU	A	233	7.228	18.018	30.549	1.00	21.20
495	CA	LEU	A	233	8.483	17.785	31.278	1.00	20.50
496	C	LEU	A	233	8.267	17.940	32.785	1.00	20.58
497	O	LEU	A	233	8.755	17.139	33.587	1.00	18.39
498	CB	LEU	A	233	9.565	18.770	30.811	1.00	20.92
502	N	VAL	A	234	7.539	18.981	33.172	1.00	20.09
503	CA	VAL	A	234	7.263	19.217	34.583	1.00	20.15
504	C	VAL	A	234	6.320	18.152	35.146	1.00	19.97
505	O	VAL	A	234	6.500	17.691	36.268	1.00	19.99
506	CB	VAL	A	234	6.665	20.630	34.796	1.00	21.02
509	N	SER	A	235	5.324	17.749	34.362	1.00	18.23
510	CA	SER	A	235	4.378	16.732	34.821	1.00	19.68
511	C	SER	A	235	5.117	15.413	35.079	1.00	19.46
512	O	SER	A	235	4.906	14.743	36.095	1.00	19.95
513	CB	SER	A	235	3.284	16.537	33.767	1.00	21.15
515	N	TYR	A	236	5.983	15.057	34.140	1.00	19.05
516	CA	TYR	A	236	6.796	13.849	34.222	1.00	19.13
517	C	TYR	A	236	7.660	13.930	35.479	1.00	18.86
518	O	TYR	A	236	7.792	12.958	36.223	1.00	18.39
519	CB	TYR	A	236	7.675	13.781	32.976	1.00	19.07
520	CG	TYR	A	236	8.800	12.764	32.990	1.00	19.18
521	CD1	TYR	A	236	8.601	11.466	32.527	1.00	20.38
522	CD2	TYR	A	236	10.084	13.131	33.391	1.00	20.55
523	CE1	TYR	A	236	9.665	10.557	32.448	1.00	21.48
524	CE2	TYR	A	236	11.149	12.233	33.321	1.00	20.66
525	CZ	TYR	A	236	10.934	10.954	32.846	1.00	21.96
526	OH	TYR	A	236	11.996	10.079	32.749	1.00	21.78

527	N	SER	A	237	8.241	15.105	35.711	1.00	17.93
528	CA	SER	A	237	9.106	15.312	36.868	1.00	18.19
529	C	SER	A	237	8.373	15.218	38.199	1.00	18.73
530	O	SER	A	237	8.929	14.737	39.184	1.00	19.34
531	CB	SER	A	237	9.830	16.654	36.730	1.00	18.72
533	N	ILE	A	238	7.128	15.680	38.237	1.00	18.89
534	CA	ILE	A	238	6.343	15.597	39.460	1.00	20.25
535	C	ILE	A	238	6.101	14.119	39.759	1.00	20.17
536	O	ILE	A	238	6.129	13.705	40.914	1.00	20.62
537	CB	ILE	A	238	4.984	16.337	39.317	1.00	21.21
541	N	GLN	A	239	5.868	13.315	38.719	1.00	20.04
542	CA	GLN	A	239	5.657	11.890	38.936	1.00	19.72
543	C	GLN	A	239	6.911	11.261	39.531	1.00	20.24
544	O	GLN	A	239	6.823	10.433	40.437	1.00	19.92
545	CB	GLN	A	239	5.288	11.178	37.628	1.00	21.35
550	N	LYS	A	240	8.080	11.661	39.037	1.00	19.37
551	CA	LYS	A	240	9.336	11.116	39.557	1.00	19.49
552	C	LYS	A	240	9.575	11.583	40.994	1.00	20.03
553	O	LYS	A	240	10.086	10.826	41.826	1.00	20.81
554	CB	LYS	A	240	10.509	11.525	38.658	1.00	19.27
559	N	VAL	A	241	9.203	12.827	41.284	1.00	19.95
560	CA	VAL	A	241	9.355	13.380	42.630	1.00	21.18
561	C	VAL	A	241	8.466	12.633	43.621	1.00	22.58
562	O	VAL	A	241	8.845	12.418	44.769	1.00	22.01
563	CB	VAL	A	241	9.006	14.890	42.658	1.00	22.53
566	N	ILE	A	242	7.277	12.237	43.178	1.00	22.44
567	CA	ILE	A	242	6.375	11.492	44.052	1.00	23.64
568	C	ILE	A	242	7.027	10.157	44.416	1.00	23.45
569	O	ILE	A	242	6.987	9.726	45.573	1.00	25.50
570	CB	ILE	A	242	5.012	11.255	43.360	1.00	24.32
571	CG1	ILE	A	242	4.235	12.575	43.303	1.00	25.64
572	CG2	ILE	A	242	4.214	10.186	44.104	1.00	24.95
573	CD1	ILE	A	242	3.012	12.540	42.401	1.00	25.41
574	N	GLY	A	243	7.652	9.521	43.431	1.00	22.76
575	CA	GLY	A	243	8.310	8.246	43.665	1.00	23.14
576	C	GLY	A	243	9.491	8.385	44.604	1.00	23.29
577	O	GLY	A	243	9.719	7.525	45.454	1.00	24.26
578	N	PHE	A	244	10.244	9.471	44.443	1.00	22.21
579	CA	PHE	A	244	11.406	9.754	45.287	1.00	23.08
580	C	PHE	A	244	10.962	9.960	46.734	1.00	23.33
581	O	PHE	A	244	11.509	9.359	47.665	1.00	22.96
582	CB	PHE	A	244	12.110	11.023	44.799	1.00	21.55
583	CG	PHE	A	244	13.264	11.454	45.663	1.00	23.20
584	CD1	PHE	A	244	14.474	10.764	45.632	1.00	25.04
585	CD2	PHE	A	244	13.140	12.548	46.516	1.00	24.78
586	CE1	PHE	A	244	15.542	11.157	46.437	1.00	25.46
587	CE2	PHE	A	244	14.205	12.950	47.327	1.00	24.71
588	CZ	PHE	A	244	15.407	12.254	47.286	1.00	24.22
589	N	ALA	A	245	9.963	10.819	46.912	1.00	23.25
590	CA	ALA	A	245	9.441	11.134	48.233	1.00	23.37
591	C	ALA	A	245	8.960	9.906	49.006	1.00	25.09
592	O	ALA	A	245	9.182	9.805	50.212	1.00	24.87
593	CB	ALA	A	245	8.310	12.156	48.113	1.00	22.36
594	N	LYS	A	246	8.309	8.975	48.314	1.00	26.15
595	CA	LYS	A	246	7.800	7.768	48.959	1.00	28.66

596	C	LYS	A	246	8.914	6.918	49.562	1.00	29.21
597	O	LYS	A	246	8.668	6.117	50.466	1.00	29.75
598	CB	LYS	A	246	6.997	6.931	47.957	1.00	30.93
599	CG	LYS	A	246	5.702	7.593	47.501	1.00	34.75
600	CD	LYS	A	246	5.017	6.811	46.383	1.00	37.28
601	CE	LYS	A	246	4.410	5.501	46.873	1.00	40.02
602	NZ	LYS	A	246	3.230	5.724	47.756	1.00	42.15
603	N	MET	A	247	10.138	7.104	49.074	1.00	28.68
604	CA	MET	A	247	11.282	6.339	49.562	1.00	29.45
605	C	MET	A	247	12.076	7.021	50.681	1.00	28.75
606	O	MET	A	247	13.012	6.431	51.230	1.00	28.61
607	CB	MET	A	247	12.219	5.990	48.396	1.00	30.97
611	N	ILE	A	248	11.709	8.253	51.023	1.00	27.31
612	CA	ILE	A	248	12.391	8.973	52.100	1.00	28.07
613	C	ILE	A	248	12.033	8.295	53.420	1.00	28.99
614	O	ILE	A	248	10.859	8.179	53.763	1.00	28.97
615	CB	ILE	A	248	11.934	10.446	52.195	1.00	27.49
616	CG1	ILE	A	248	12.299	11.198	50.916	1.00	26.51
617	CG2	ILE	A	248	12.582	11.114	53.411	1.00	28.63
618	CD1	ILE	A	248	11.730	12.605	50.865	1.00	25.47
619	N	PRO	A	249	13.041	7.844	54.181	1.00	30.64
620	CA	PRO	A	249	12.764	7.182	55.460	1.00	31.91
621	C	PRO	A	249	11.818	7.992	56.348	1.00	32.16
622	O	PRO	A	249	12.107	9.138	56.688	1.00	34.02
623	CB	PRO	A	249	14.153	7.035	56.075	1.00	32.36
626	N	GLY	A	250	10.686	7.392	56.706	1.00	32.30
627	CA	GLY	A	250	9.725	8.064	57.565	1.00	33.31
628	C	GLY	A	250	8.542	8.700	56.858	1.00	33.48
629	O	GLY	A	250	7.484	8.888	57.459	1.00	33.45
630	N	PHE	A	251	8.709	9.023	55.579	1.00	33.51
631	CA	PHE	A	251	7.643	9.658	54.809	1.00	33.69
632	C	PHE	A	251	6.335	8.871	54.833	1.00	34.57
633	O	PHE	A	251	5.259	9.455	54.964	1.00	35.10
634	CB	PHE	A	251	8.082	9.850	53.356	1.00	31.35
635	CG	PHE	A	251	7.180	10.754	52.564	1.00	29.89
636	CD1	PHE	A	251	7.234	12.134	52.735	1.00	30.12
637	CD2	PHE	A	251	6.276	10.227	51.643	1.00	30.05
638	CE1	PHE	A	251	6.400	12.979	51.999	1.00	29.28
639	CE2	PHE	A	251	5.441	11.063	50.906	1.00	28.78
640	CZ	PHE	A	251	5.505	12.440	51.085	1.00	28.48
641	N	ARG	A	252	6.431	7.551	54.703	1.00	36.83
642	CA	ARG	A	252	5.250	6.691	54.698	1.00	39.19
643	C	ARG	A	252	4.535	6.647	56.045	1.00	39.61
644	O	ARG	A	252	3.391	6.200	56.127	1.00	40.31
645	CB	ARG	A	252	5.625	5.262	54.292	1.00	41.06
646	CG	ARG	A	252	6.138	5.101	52.867	1.00	44.96
647	CD	ARG	A	252	6.260	3.620	52.516	1.00	47.63
648	NE	ARG	A	252	6.777	3.393	51.169	1.00	50.79
649	CZ	ARG	A	252	8.062	3.459	50.831	1.00	51.79
650	NH1	ARG	A	252	8.982	3.745	51.745	1.00	52.82
651	NH2	ARG	A	252	8.427	3.235	49.576	1.00	52.64
652	N	ASP	A	253	5.205	7.102	57.098	1.00	39.96
653	CA	ASP	A	253	4.610	7.097	58.430	1.00	40.45
654	C	ASP	A	253	3.648	8.255	58.635	1.00	39.90
655	O	ASP	A	253	2.902	8.284	59.612	1.00	39.68

656	CB	ASP	A	253	5.698	7.127	59.506	1.00	42.53
657	CG	ASP	A	253	6.524	5.856	59.531	1.00	44.84
658	OD1	ASP	A	253	5.938	4.767	59.345	1.00	47.60
659	OD2	ASP	A	253	7.752	5.942	59.743	1.00	45.66
660	N	LEU	A	254	3.669	9.208	57.710	1.00	38.00
661	CA	LEU	A	254	2.782	10.361	57.780	1.00	37.81
662	C	LEU	A	254	1.417	9.978	57.218	1.00	37.76
663	O	LEU	A	254	1.293	9.000	56.476	1.00	37.49
664	CB	LEU	A	254	3.348	11.521	56.955	1.00	36.51
665	CG	LEU	A	254	4.707	12.101	57.346	1.00	36.86
666	CD1	LEU	A	254	5.142	13.113	56.297	1.00	35.66
667	CD2	LEU	A	254	4.620	12.751	58.719	1.00	36.85
668	N	THR	A	255	0.395	10.745	57.579	1.00	38.14
669	CA	THR	A	255	-0.950	10.496	57.077	1.00	39.08
670	C	THR	A	255	-0.928	10.828	55.588	1.00	39.76
671	O	THR	A	255	-0.075	11.593	55.139	1.00	39.15
672	CB	THR	A	255	-1.982	11.397	57.781	1.00	39.68
675	N	SER	A	256	-1.851	10.258	54.817	1.00	40.24
676	CA	SER	A	256	-1.883	10.527	53.382	1.00	40.92
677	C	SER	A	256	-2.148	12.010	53.152	1.00	40.23
678	O	SER	A	256	-1.662	12.599	52.185	1.00	40.28
679	CB	SER	A	256	-2.968	9.693	52.690	1.00	41.53
681	N	GLU	A	257	-2.916	12.610	54.056	1.00	39.55
682	CA	GLU	A	257	-3.252	14.024	53.963	1.00	38.71
683	C	GLU	A	257	-1.999	14.889	54.038	1.00	36.50
684	O	GLU	A	257	-1.825	15.810	53.240	1.00	36.20
685	CB	GLU	A	257	-4.221	14.400	55.085	1.00	41.43
686	CG	GLU	A	257	-4.650	15.853	55.090	1.00	44.63
687	CD	GLU	A	257	-5.747	16.121	56.103	1.00	47.39
688	OE1	GLU	A	257	-6.879	15.634	55.896	1.00	48.76
689	OE2	GLU	A	257	-5.476	16.810	57.109	1.00	48.99
690	N	ASP	A	258	-1.132	14.593	55.001	1.00	34.59
691	CA	ASP	A	258	0.111	15.339	55.159	1.00	33.04
692	C	ASP	A	258	1.064	15.047	54.002	1.00	32.48
693	O	ASP	A	258	1.782	15.934	53.546	1.00	31.37
694	CB	ASP	A	258	0.784	14.984	56.488	1.00	34.07
695	CG	ASP	A	258	0.256	15.809	57.645	1.00	35.11
696	OD1	ASP	A	258	0.599	15.501	58.807	1.00	35.63
697	OD2	ASP	A	258	-0.493	16.775	57.386	1.00	34.86
698	N	GLN	A	259	1.072	13.803	53.532	1.00	31.90
699	CA	GLN	A	259	1.940	13.433	52.417	1.00	32.81
700	C	GLN	A	259	1.611	14.272	51.184	1.00	32.59
701	O	GLN	A	259	2.505	14.820	50.534	1.00	32.51
702	CB	GLN	A	259	1.783	11.946	52.077	1.00	32.98
703	CG	GLN	A	259	2.217	11.000	53.181	1.00	34.94
704	CD	GLN	A	259	2.168	9.547	52.755	1.00	37.19
705	OE1	GLN	A	259	2.322	8.641	53.576	1.00	39.55
706	NE2	GLN	A	259	1.958	9.315	51.466	1.00	37.81
707	N	ILE	A	260	0.325	14.375	50.866	1.00	32.68
708	CA	ILE	A	260	-0.109	15.147	49.706	1.00	32.42
709	C	ILE	A	260	0.183	16.634	49.880	1.00	31.57
710	O	ILE	A	260	0.588	17.311	48.933	1.00	30.43
711	CB	ILE	A	260	-1.619	14.959	49.445	1.00	33.97
712	CG1	ILE	A	260	-1.933	13.471	49.277	1.00	34.59
713	CG2	ILE	A	260	-2.036	15.731	48.201	1.00	33.83

714	CD1	ILE	A	260	-1.156	12.789	48.165	1.00	36.85
715	N	VAL	A	261	-0.029	17.146	51.088	1.00	29.87
716	CA	VAL	A	261	0.244	18.551	51.358	1.00	28.91
717	C	VAL	A	261	1.717	18.862	51.097	1.00	27.91
718	O	VAL	A	261	2.043	19.856	50.445	1.00	28.13
719	CB	VAL	A	261	-0.089	18.923	52.827	1.00	28.91
722	N	LEU	A	262	2.605	18.011	51.604	1.00	27.17
723	CA	LEU	A	262	4.039	18.222	51.423	1.00	25.81
724	C	LEU	A	262	4.461	18.126	49.955	1.00	25.46
725	O	LEU	A	262	5.274	18.921	49.485	1.00	24.78
726	CB	LEU	A	262	4.836	17.219	52.265	1.00	26.02
727	CG	LEU	A	262	4.604	17.278	53.781	1.00	25.71
728	CD1	LEU	A	262	5.382	16.162	54.464	1.00	26.98
729	CD2	LEU	A	262	5.028	18.634	54.317	1.00	26.30
730	N	LEU	A	263	3.911	17.155	49.232	1.00	25.71
731	CA	LEU	A	263	4.244	16.985	47.818	1.00	26.15
732	C	LEU	A	263	3.763	18.166	46.974	1.00	26.42
733	O	LEU	A	263	4.514	18.702	46.154	1.00	25.91
734	CB	LEU	A	263	3.633	15.681	47.283	1.00	27.20
735	CG	LEU	A	263	4.293	14.376	47.745	1.00	29.34
736	CD1	LEU	A	263	3.401	13.197	47.404	1.00	30.14
737	CD2	LEU	A	263	5.658	14.223	47.082	1.00	31.00
738	N	LYS	A	264	2.519	18.585	47.178	1.00	25.69
739	CA	LYS	A	264	1.987	19.699	46.405	1.00	26.59
740	C	LYS	A	264	2.709	21.011	46.655	1.00	26.97
741	O	LYS	A	264	2.962	21.767	45.723	1.00	27.99
742	CB	LYS	A	264	0.496	19.899	46.688	1.00	29.36
743	CG	LYS	A	264	-0.417	18.910	45.994	1.00	31.84
744	CD	LYS	A	264	-1.862	19.348	46.156	1.00	35.05
745	CE	LYS	A	264	-2.822	18.400	45.468	1.00	37.95
746	NZ	LYS	A	264	-4.233	18.872	45.629	1.00	39.41
747	N	SER	A	265	3.047	21.286	47.908	1.00	26.64
748	CA	SER	A	265	3.712	22.540	48.227	1.00	27.75
749	C	SER	A	265	5.199	22.591	47.884	1.00	26.92
750	O	SER	A	265	5.750	23.676	47.723	1.00	28.28
751	CB	SER	A	265	3.513	22.881	49.709	1.00	28.81
753	N	SER	A	266	5.847	21.434	47.757	1.00	25.17
754	CA	SER	A	266	7.277	21.412	47.449	1.00	23.91
755	C	SER	A	266	7.609	21.035	46.011	1.00	23.51
756	O	SER	A	266	8.749	21.206	45.572	1.00	23.30
757	CB	SER	A	266	8.001	20.445	48.385	1.00	24.45
759	N	ALA	A	267	6.619	20.519	45.285	1.00	22.67
760	CA	ALA	A	267	6.801	20.089	43.898	1.00	23.39
761	C	ALA	A	267	7.698	20.979	43.040	1.00	23.51
762	O	ALA	A	267	8.716	20.517	42.515	1.00	23.55
763	CB	ALA	A	267	5.436	19.938	43.217	1.00	24.51
764	N	ILE	A	268	7.330	22.247	42.883	1.00	22.01
765	CA	ILE	A	268	8.132	23.135	42.041	1.00	22.23
766	C	ILE	A	268	9.539	23.374	42.592	1.00	22.05
767	O	ILE	A	268	10.494	23.558	41.828	1.00	20.90
768	CB	ILE	A	268	7.426	24.496	41.811	1.00	23.63
772	N	GLU	A	269	9.674	23.352	43.911	1.00	20.40
773	CA	GLU	A	269	10.979	23.561	44.529	1.00	20.63
774	C	GLU	A	269	11.933	22.402	44.268	1.00	21.33
775	O	GLU	A	269	13.109	22.620	43.976	1.00	20.99

776	CB	GLU	A	269	10.823	23.770	46.030	1.00	20.38
777	CG	GLU	A	269	10.206	25.110	46.396	1.00	22.10
778	CD	GLU	A	269	10.009	25.261	47.892	1.00	23.72
779	OE1	GLU	A	269	10.803	24.670	48.656	1.00	22.73
780	OE2	GLU	A	269	9.067	25.974	48.301	1.00	24.58
781	N	VAL	A	270	11.434	21.172	44.375	1.00	20.39
782	CA	VAL	A	270	12.279	20.006	44.143	1.00	20.83
783	C	VAL	A	270	12.644	19.911	42.670	1.00	20.52
784	O	VAL	A	270	13.734	19.458	42.318	1.00	20.87
785	CB	VAL	A	270	11.582	18.709	44.597	1.00	21.55
788	N	ILE	A	271	11.731	20.337	41.804	1.00	20.29
789	CA	ILE	A	271	12.010	20.318	40.376	1.00	20.71
790	C	ILE	A	271	13.145	21.300	40.099	1.00	20.86
791	O	ILE	A	271	14.083	20.990	39.361	1.00	20.78
792	CB	ILE	A	271	10.755	20.684	39.563	1.00	21.89
796	N	MET	A	272	13.076	22.481	40.701	1.00	21.17
797	CA	MET	A	272	14.147	23.446	40.500	1.00	21.57
798	C	MET	A	272	15.474	22.888	41.020	1.00	20.82
799	O	MET	A	272	16.513	23.064	40.384	1.00	22.20
800	CB	MET	A	272	13.800	24.770	41.183	1.00	22.31
804	N	LEU	A	273	15.442	22.204	42.163	1.00	21.17
805	CA	LEU	A	273	16.661	21.606	42.717	1.00	21.28
806	C	LEU	A	273	17.226	20.486	41.842	1.00	20.96
807	O	LEU	A	273	18.408	20.494	41.487	1.00	20.75
808	CB	LEU	A	273	16.405	21.026	44.116	1.00	22.98
809	CG	LEU	A	273	16.367	21.940	45.337	1.00	25.62
810	CD1	LEU	A	273	15.959	21.129	46.572	1.00	25.83
811	CD2	LEU	A	273	17.736	22.571	45.543	1.00	26.65
812	N	ARG	A	274	16.385	19.517	41.494	1.00	19.69
813	CA	ARG	A	274	16.852	18.384	40.702	1.00	19.52
814	C	ARG	A	274	17.317	18.787	39.309	1.00	19.10
815	O	ARG	A	274	18.159	18.117	38.715	1.00	19.83
816	CB	ARG	A	274	15.759	17.299	40.610	1.00	19.75
817	CG	ARG	A	274	14.652	17.566	39.601	1.00	19.52
818	CD	ARG	A	274	13.381	16.792	39.969	1.00	19.72
819	NE	ARG	A	274	13.599	15.356	40.153	1.00	18.11
820	CZ	ARG	A	274	13.580	14.453	39.175	1.00	19.01
821	NH1	ARG	A	274	13.357	14.824	37.919	1.00	18.53
822	NH2	ARG	A	274	13.759	13.168	39.458	1.00	19.51
823	N	SER	A	275	16.792	19.892	38.793	1.00	19.73
824	CA	SER	A	275	17.183	20.331	37.463	1.00	19.93
825	C	SER	A	275	18.615	20.838	37.442	1.00	19.90
826	O	SER	A	275	19.191	21.016	36.377	1.00	20.21
827	CB	SER	A	275	16.249	21.437	36.958	1.00	20.51
829	N	ASN	A	276	19.198	21.055	38.615	1.00	20.28
830	CA	ASN	A	276	20.564	21.557	38.662	1.00	19.85
831	C	ASN	A	276	21.512	20.544	38.024	1.00	21.26
832	O	ASN	A	276	22.585	20.903	37.538	1.00	19.72
833	CB	ASN	A	276	20.983	21.843	40.108	1.00	20.77
834	CG	ASN	A	276	22.265	22.651	40.187	1.00	23.39
835	OD1	ASN	A	276	23.275	22.187	40.713	1.00	26.18
836	ND2	ASN	A	276	22.231	23.867	39.649	1.00	21.92
837	N	GLU	A	277	21.096	19.280	38.000	1.00	20.52
838	CA	GLU	A	277	21.925	18.226	37.425	1.00	21.75
839	C	GLU	A	277	22.103	18.370	35.908	1.00	21.79

840	O	GLU	A	277	23.105	17.910	35.351	1.00	22.41
841	CB	GLU	A	277	21.331	16.852	37.785	1.00	22.91
846	N	SER	A	278	21.152	19.011	35.234	1.00	19.68
847	CA	SER	A	278	21.266	19.194	33.789	1.00	20.64
848	C	SER	A	278	21.712	20.607	33.448	1.00	21.58
849	O	SER	A	278	22.008	20.910	32.292	1.00	22.05
850	CB	SER	A	278	19.934	18.910	33.092	1.00	20.93
852	N	PHE	A	279	21.751	21.474	34.451	1.00	21.92
853	CA	PHE	A	279	22.160	22.853	34.219	1.00	23.24
854	C	PHE	A	279	23.659	22.912	33.972	1.00	24.55
855	O	PHE	A	279	24.429	22.218	34.638	1.00	24.49
856	CB	PHE	A	279	21.820	23.723	35.429	1.00	23.08
857	CG	PHE	A	279	22.051	25.187	35.198	1.00	24.02
858	CD1	PHE	A	279	21.135	25.942	34.471	1.00	24.96
859	CD2	PHE	A	279	23.197	25.805	35.682	1.00	24.94
860	CE1	PHE	A	279	21.356	27.293	34.227	1.00	24.93
861	CE2	PHE	A	279	23.429	27.160	35.442	1.00	25.50
862	CZ	PHE	A	279	22.506	27.903	34.714	1.00	24.47
863	N	THR	A	280	24.077	23.728	33.010	1.00	24.73
864	CA	THR	A	280	25.496	23.872	32.728	1.00	26.87
865	C	THR	A	280	25.884	25.343	32.672	1.00	27.44
866	O	THR	A	280	25.186	26.162	32.070	1.00	26.28
867	CB	THR	A	280	25.897	23.198	31.399	1.00	27.76
870	N	MET	A	281	26.991	25.676	33.326	1.00	28.33
871	CA	MET	A	281	27.469	27.049	33.340	1.00	31.03
872	C	MET	A	281	28.275	27.390	32.095	1.00	31.28
873	O	MET	A	281	28.812	28.490	31.980	1.00	30.87
874	CB	MET	A	281	28.298	27.306	34.596	1.00	33.43
878	N	ASP	A	282	28.364	26.448	31.159	1.00	31.72
879	CA	ASP	A	282	29.097	26.709	29.925	1.00	32.91
880	C	ASP	A	282	28.366	27.818	29.175	1.00	32.02
881	O	ASP	A	282	28.989	28.764	28.683	1.00	31.15
882	CB	ASP	A	282	29.172	25.455	29.050	1.00	35.93
886	N	ASP	A	283	27.041	27.702	29.100	1.00	29.87
887	CA	ASP	A	283	26.224	28.704	28.418	1.00	28.59
888	C	ASP	A	283	24.931	29.032	29.170	1.00	27.92
889	O	ASP	A	283	23.984	29.568	28.592	1.00	27.21
890	CB	ASP	A	283	25.904	28.243	26.994	1.00	29.84
894	N	MET	A	284	24.902	28.708	30.460	1.00	26.84
895	CA	MET	A	284	23.748	28.985	31.317	1.00	27.62
896	C	MET	A	284	22.449	28.379	30.801	1.00	27.20
897	O	MET	A	284	21.429	29.060	30.686	1.00	27.86
898	CB	MET	A	284	23.565	30.497	31.484	1.00	29.95
902	N	SER	A	285	22.479	27.091	30.503	1.00	25.58
903	CA	SER	A	285	21.288	26.427	30.010	1.00	24.62
904	C	SER	A	285	21.136	25.090	30.697	1.00	24.72
905	O	SER	A	285	22.028	24.641	31.415	1.00	24.18
906	CB	SER	A	285	21.402	26.186	28.509	1.00	24.98
908	N	TRP	A	286	19.982	24.472	30.480	1.00	24.17
909	CA	TRP	A	286	19.699	23.146	30.997	1.00	24.74
910	C	TRP	A	286	19.842	22.312	29.732	1.00	25.34
911	O	TRP	A	286	19.006	22.391	28.828	1.00	25.37
912	CB	TRP	A	286	18.268	23.064	31.522	1.00	23.76
913	CG	TRP	A	286	18.048	23.702	32.863	1.00	21.76
914	CD1	TRP	A	286	18.186	23.107	34.088	1.00	21.47

915	CD2	TRP	A	286	17.568	25.031	33.118	1.00	23.03
916	NE1	TRP	A	286	17.811	23.976	35.084	1.00	21.88
917	CE2	TRP	A	286	17.429	25.164	34.519	1.00	22.96
918	CE3	TRP	A	286	17.238	26.121	32.299	1.00	23.54
919	CZ2	TRP	A	286	16.970	26.341	35.120	1.00	24.15
920	CZ3	TRP	A	286	16.781	27.293	32.898	1.00	22.92
921	CH2	TRP	A	286	16.651	27.390	34.297	1.00	23.66
922	N	THR	A	287	20.918	21.540	29.654	1.00	25.53
923	CA	THR	A	287	21.173	20.721	28.478	1.00	27.14
924	C	THR	A	287	20.833	19.266	28.753	1.00	27.53
925	O	THR	A	287	21.501	18.607	29.551	1.00	27.40
926	CB	THR	A	287	22.644	20.853	28.049	1.00	27.77
929	N	CYS	A	288	19.792	18.775	28.084	1.00	28.08
930	CA	CYS	A	288	19.326	17.406	28.270	1.00	30.34
931	C	CYS	A	288	19.478	16.520	27.040	1.00	33.66
932	O	CYS	A	288	18.530	15.857	26.624	1.00	33.19
933	CB	CYS	A	288	17.861	17.426	28.699	1.00	29.32
934	SG	CYS	A	288	17.566	18.403	30.188	1.00	28.01
935	N	GLY	A	289	20.675	16.498	26.466	1.00	37.69
936	CA	GLY	A	289	20.897	15.682	25.286	1.00	41.85
937	C	GLY	A	289	21.072	16.536	24.044	1.00	44.11
938	O	GLY	A	289	21.842	17.497	24.051	1.00	45.10
939	N	ASN	A	290	20.349	16.205	22.978	1.00	46.33
940	CA	ASN	A	290	20.469	16.959	21.737	1.00	47.32
941	C	ASN	A	290	19.961	18.391	21.874	1.00	47.22
942	O	ASN	A	290	19.303	18.746	22.857	1.00	47.49
943	CB	ASN	A	290	19.733	16.241	20.600	1.00	49.56
947	N	GLN	A	291	20.277	19.205	20.874	1.00	46.24
948	CA	GLN	A	291	19.896	20.611	20.850	1.00	45.60
949	C	GLN	A	291	18.402	20.859	21.031	1.00	43.20
950	O	GLN	A	291	18.007	21.916	21.520	1.00	43.27
951	CB	GLN	A	291	20.380	21.247	19.545	1.00	47.46
956	N	ASP	A	292	17.574	19.897	20.636	1.00	40.92
957	CA	ASP	A	292	16.129	20.046	20.780	1.00	38.58
958	C	ASP	A	292	15.740	20.140	22.252	1.00	35.80
959	O	ASP	A	292	14.769	20.814	22.601	1.00	34.04
960	CB	ASP	A	292	15.391	18.862	20.145	1.00	41.69
964	N	TYR	A	293	16.506	19.469	23.111	1.00	33.09
965	CA	TYR	A	293	16.219	19.465	24.543	1.00	31.43
966	C	TYR	A	293	17.183	20.305	25.367	1.00	30.28
967	O	TYR	A	293	17.558	19.934	26.481	1.00	30.56
968	CB	TYR	A	293	16.186	18.027	25.066	1.00	31.64
969	CG	TYR	A	293	15.232	17.154	24.287	1.00	31.43
970	CD1	TYR	A	293	15.591	15.864	23.905	1.00	32.22
971	CD2	TYR	A	293	13.999	17.647	23.861	1.00	32.09
972	CE1	TYR	A	293	14.752	15.091	23.106	1.00	32.84
973	CE2	TYR	A	293	13.153	16.883	23.063	1.00	31.78
974	CZ	TYR	A	293	13.537	15.611	22.684	1.00	33.14
975	OH	TYR	A	293	12.726	14.874	21.850	1.00	32.75
1008	N	VAL	A	297	16.879	30.678	28.634	1.00	29.11
1009	CA	VAL	A	297	15.956	31.675	29.167	1.00	30.41
1010	C	VAL	A	297	14.821	31.972	28.187	1.00	30.24
1011	O	VAL	A	297	13.655	32.065	28.582	1.00	29.94
1012	CB	VAL	A	297	16.692	33.005	29.475	1.00	30.71
1015	N	SER	A	298	15.168	32.115	26.912	1.00	30.44

1016	CA	SER	A	298	14.185	32.430	25.881	1.00	30.65
1017	C	SER	A	298	13.106	31.370	25.714	1.00	30.99
1018	O	SER	A	298	11.986	31.680	25.304	1.00	31.34
1019	CB	SER	A	298	14.884	32.675	24.539	1.00	31.86
1021	N	ASP	A	299	13.435	30.121	26.028	1.00	29.88
1022	CA	ASP	A	299	12.464	29.042	25.912	1.00	29.41
1023	C	ASP	A	299	11.424	29.137	27.019	1.00	28.20
1024	O	ASP	A	299	10.268	28.770	26.827	1.00	28.75
1025	CB	ASP	A	299	13.162	27.679	25.979	1.00	31.09
1029	N	VAL	A	300	11.837	29.631	28.183	1.00	27.61
1030	CA	VAL	A	300	10.923	29.760	29.308	1.00	26.53
1031	C	VAL	A	300	9.948	30.913	29.070	1.00	26.97
1032	O	VAL	A	300	8.781	30.835	29.449	1.00	26.32
1033	CB	VAL	A	300	11.703	29.972	30.623	1.00	27.74
1036	N	THR	A	301	10.420	31.980	28.432	1.00	26.55
1037	CA	THR	A	301	9.539	33.106	28.142	1.00	27.35
1038	C	THR	A	301	8.507	32.672	27.100	1.00	27.20
1039	O	THR	A	301	7.394	33.188	27.069	1.00	27.90
1040	CB	THR	A	301	10.324	34.329	27.617	1.00	27.90
1041	OG1	THR	A	301	11.097	33.956	26.472	1.00	29.74
1042	CG2	THR	A	301	11.250	34.861	28.696	1.00	29.44
1043	N	LYS	A	302	8.875	31.715	26.250	1.00	26.49
1044	CA	LYS	A	302	7.948	31.225	25.232	1.00	27.28
1045	C	LYS	A	302	6.886	30.318	25.847	1.00	27.81
1046	O	LYS	A	302	5.960	29.874	25.160	1.00	27.95
1047	CB	LYS	A	302	8.701	30.477	24.130	1.00	28.36
1052	N	ALA	A	303	7.019	30.048	27.143	1.00	26.44
1053	CA	ALA	A	303	6.052	29.219	27.847	1.00	27.88
1054	C	ALA	A	303	5.130	30.097	28.692	1.00	28.91
1055	O	ALA	A	303	4.310	29.592	29.457	1.00	29.81
1056	CB	ALA	A	303	6.771	28.199	28.726	1.00	27.38
1057	N	GLY	A	304	5.279	31.415	28.564	1.00	29.66
1058	CA	GLY	A	304	4.423	32.328	29.309	1.00	30.57
1059	C	GLY	A	304	4.963	32.961	30.582	1.00	31.32
1060	O	GLY	A	304	4.257	33.735	31.234	1.00	32.07
1061	N	HIS	A	305	6.202	32.649	30.948	1.00	31.10
1062	CA	HIS	A	305	6.797	33.216	32.155	1.00	30.95
1063	C	HIS	A	305	7.656	34.439	31.853	1.00	31.77
1064	O	HIS	A	305	8.138	34.610	30.731	1.00	31.65
1065	CB	HIS	A	305	7.628	32.155	32.881	1.00	30.92
1071	N	SER	A	306	7.839	35.290	32.860	1.00	32.01
1072	CA	SER	A	306	8.624	36.511	32.700	1.00	33.97
1073	C	SER	A	306	9.982	36.449	33.392	1.00	34.00
1074	O	SER	A	306	10.265	35.523	34.154	1.00	33.09
1075	CB	SER	A	306	7.842	37.710	33.235	1.00	34.42
1077	N	LEU	A	307	10.813	37.455	33.125	1.00	34.07
1078	CA	LEU	A	307	12.155	37.537	33.694	1.00	34.93
1079	C	LEU	A	307	12.172	37.666	35.212	1.00	33.80
1080	O	LEU	A	307	13.180	37.364	35.851	1.00	33.69
1081	CB	LEU	A	307	12.923	38.710	33.068	1.00	36.84
1085	N	GLU	A	308	11.060	38.110	35.789	1.00	33.38
1086	CA	GLU	A	308	10.963	38.265	37.235	1.00	32.81
1087	C	GLU	A	308	11.165	36.913	37.917	1.00	31.88
1088	O	GLU	A	308	11.558	36.842	39.078	1.00	30.22
1089	CB	GLU	A	308	9.603	38.856	37.607	1.00	37.03

1094	N	LEU	A	309	10.898	35.838	37.182	1.00	29.69
1095	CA	LEU	A	309	11.081	34.492	37.714	1.00	29.34
1096	C	LEU	A	309	12.348	33.872	37.130	1.00	28.31
1097	O	LEU	A	309	13.160	33.290	37.848	1.00	26.92
1098	CB	LEU	A	309	9.882	33.605	37.360	1.00	28.48
1099	CG	LEU	A	309	10.037	32.116	37.700	1.00	28.85
1100	CD1	LEU	A	309	10.011	31.931	39.211	1.00	29.55
1101	CD2	LEU	A	309	8.919	31.312	37.048	1.00	29.07
1102	N	ILE	A	310	12.524	34.019	35.822	1.00	28.87
1103	CA	ILE	A	310	13.673	33.428	35.142	1.00	30.36
1104	C	ILE	A	310	15.051	33.907	35.590	1.00	30.97
1105	O	ILE	A	310	15.948	33.092	35.808	1.00	30.03
1106	CB	ILE	A	310	13.552	33.605	33.617	1.00	31.31
1110	N	GLU	A	311	15.240	35.213	35.726	1.00	31.83
1111	CA	GLU	A	311	16.547	35.707	36.151	1.00	33.28
1112	C	GLU	A	311	16.945	35.175	37.528	1.00	31.76
1113	O	GLU	A	311	18.067	34.707	37.714	1.00	31.24
1114	CB	GLU	A	311	16.573	37.237	36.128	1.00	35.65
1115	CG	GLU	A	311	16.550	37.788	34.710	1.00	41.13
1116	CD	GLU	A	311	16.753	39.287	34.649	1.00	43.32
1117	OE1	GLU	A	311	16.858	39.815	33.522	1.00	46.68
1118	OE2	GLU	A	311	16.807	39.933	35.718	1.00	45.68
1119	N	PRO	A	312	16.032	35.232	38.511	1.00	30.94
1120	CA	PRO	A	312	16.358	34.728	39.851	1.00	29.89
1121	C	PRO	A	312	16.570	33.212	39.817	1.00	28.28
1122	O	PRO	A	312	17.321	32.656	40.619	1.00	28.14
1123	CB	PRO	A	312	15.132	35.115	40.675	1.00	30.62
1124	CG	PRO	A	312	14.612	36.330	39.962	1.00	31.93
1125	CD	PRO	A	312	14.740	35.943	38.523	1.00	31.29
1126	N	LEU	A	313	15.896	32.550	38.883	1.00	26.85
1127	CA	LEU	A	313	16.013	31.102	38.739	1.00	26.51
1128	C	LEU	A	313	17.425	30.764	38.267	1.00	25.16
1129	O	LEU	A	313	18.063	29.855	38.788	1.00	24.33
1130	CB	LEU	A	313	14.998	30.583	37.715	1.00	27.97
1134	N	ILE	A	314	17.917	31.504	37.279	1.00	25.12
1135	CA	ILE	A	314	19.262	31.255	36.763	1.00	25.36
1136	C	ILE	A	314	20.304	31.552	37.839	1.00	25.44
1137	O	ILE	A	314	21.267	30.802	38.008	1.00	25.07
1138	CB	ILE	A	314	19.565	32.122	35.517	1.00	26.51
1142	N	LYS	A	315	20.112	32.641	38.574	1.00	25.44
1143	CA	LYS	A	315	21.058	32.994	39.626	1.00	26.52
1144	C	LYS	A	315	21.117	31.869	40.656	1.00	25.66
1145	O	LYS	A	315	22.193	31.522	41.149	1.00	25.67
1146	CB	LYS	A	315	20.651	34.310	40.296	1.00	28.77
1147	CG	LYS	A	315	21.759	34.926	41.134	1.00	34.86
1148	CD	LYS	A	315	21.562	36.427	41.306	1.00	37.34
1149	CE	LYS	A	315	22.806	37.082	41.891	1.00	39.12
1150	NZ	LYS	A	315	23.154	36.521	43.227	1.00	41.56
1151	N	PHE	A	316	19.958	31.295	40.967	1.00	23.92
1152	CA	PHE	A	316	19.874	30.196	41.921	1.00	23.22
1153	C	PHE	A	316	20.662	28.997	41.400	1.00	22.36
1154	O	PHE	A	316	21.422	28.380	42.151	1.00	22.35
1155	CB	PHE	A	316	18.410	29.791	42.144	1.00	24.22
1156	CG	PHE	A	316	18.242	28.546	42.979	1.00	26.30
1157	CD1	PHE	A	316	18.323	28.605	44.370	1.00	27.43

1158	CD2	PHE	A	316	18.037	27.310	42.372	1.00	26.87
1159	CE1	PHE	A	316	18.204	27.446	45.141	1.00	28.46
1160	CE2	PHE	A	316	17.918	26.145	43.135	1.00	27.51
1161	CZ	PHE	A	316	18.002	26.218	44.520	1.00	28.27
1162	N	GLN	A	317	20.480	28.665	40.120	1.00	21.28
1163	CA	GLN	A	317	21.175	27.524	39.522	1.00	21.35
1164	C	GLN	A	317	22.694	27.681	39.586	1.00	21.92
1165	O	GLN	A	317	23.410	26.735	39.913	1.00	20.68
1166	CB	GLN	A	317	20.754	27.324	38.057	1.00	21.98
1171	N	VAL	A	318	23.188	28.870	39.259	1.00	22.58
1172	CA	VAL	A	318	24.629	29.108	39.301	1.00	23.76
1173	C	VAL	A	318	25.162	28.983	40.734	1.00	24.71
1174	O	VAL	A	318	26.199	28.349	40.971	1.00	26.38
1175	CB	VAL	A	318	24.975	30.510	38.727	1.00	24.56
1178	N	GLY	A	319	24.447	29.574	41.687	1.00	25.34
1179	CA	GLY	A	319	24.868	29.515	43.076	1.00	26.42
1180	C	GLY	A	319	24.892	28.099	43.623	1.00	26.70
1181	O	GLY	A	319	25.778	27.738	44.399	1.00	26.15
1182	N	LEU	A	320	23.915	27.292	43.226	1.00	25.08
1183	CA	LEU	A	320	23.856	25.910	43.680	1.00	26.49
1184	C	LEU	A	320	25.001	25.141	43.019	1.00	26.16
1185	O	LEU	A	320	25.674	24.342	43.666	1.00	25.62
1186	CB	LEU	A	320	22.499	25.289	43.318	1.00	26.17
1187	CG	LEU	A	320	22.202	23.895	43.877	1.00	29.00
1188	CD1	LEU	A	320	22.305	23.911	45.394	1.00	28.44
1189	CD2	LEU	A	320	20.803	23.457	43.439	1.00	27.28
1190	N	LYS	A	321	25.231	25.402	41.734	1.00	26.97
1191	CA	LYS	A	321	26.312	24.743	41.000	1.00	29.33
1192	C	LYS	A	321	27.664	24.983	41.649	1.00	30.36
1193	O	LYS	A	321	28.486	24.070	41.746	1.00	30.00
1194	CB	LYS	A	321	26.385	25.252	39.561	1.00	30.33
1195	CG	LYS	A	321	25.578	24.465	38.559	1.00	33.36
1196	CD	LYS	A	321	26.140	23.069	38.341	1.00	33.34
1197	CE	LYS	A	321	25.279	22.329	37.336	1.00	33.36
1198	NZ	LYS	A	321	25.668	20.911	37.111	1.00	32.77
1199	N	LYS	A	322	27.894	26.222	42.077	1.00	30.82
1200	CA	LYS	A	322	29.155	26.601	42.702	1.00	32.19
1201	C	LYS	A	322	29.447	25.934	44.037	1.00	32.03
1202	O	LYS	A	322	30.598	25.896	44.462	1.00	32.87
1203	CB	LYS	A	322	29.234	28.122	42.866	1.00	33.78
1204	CG	LYS	A	322	29.592	28.853	41.587	1.00	37.24
1205	CD	LYS	A	322	29.849	30.328	41.856	1.00	39.61
1206	CE	LYS	A	322	30.611	30.964	40.712	1.00	41.25
1207	NZ	LYS	A	322	31.956	30.335	40.544	1.00	43.80
1208	N	LEU	A	323	28.420	25.415	44.703	1.00	30.51
1209	CA	LEU	A	323	28.627	24.747	45.985	1.00	31.09
1210	C	LEU	A	323	29.296	23.392	45.774	1.00	31.05
1211	O	LEU	A	323	29.833	22.805	46.715	1.00	31.05
1212	CB	LEU	A	323	27.297	24.544	46.719	1.00	30.29
1213	CG	LEU	A	323	26.551	25.784	47.220	1.00	31.62
1214	CD1	LEU	A	323	25.260	25.359	47.904	1.00	30.41
1215	CD2	LEU	A	323	27.434	26.570	48.180	1.00	31.32
1216	N	ASN	A	324	29.264	22.908	44.535	1.00	30.91
1217	CA	ASN	A	324	29.854	21.619	44.180	1.00	32.42
1218	C	ASN	A	324	29.466	20.524	45.165	1.00	32.07

1219	O	ASN	A	324	30.323	19.864	45.755	1.00	32.62
1220	CB	ASN	A	324	31.380	21.722	44.110	1.00	36.14
1224	N	LEU	A	325	28.166	20.326	45.333	1.00	29.80
1225	CA	LEU	A	325	27.667	19.320	46.257	1.00	27.98
1226	C	LEU	A	325	27.969	17.890	45.836	1.00	27.42
1227	O	LEU	A	325	27.984	17.568	44.648	1.00	27.50
1228	CB	LEU	A	325	26.149	19.454	46.409	1.00	28.15
1229	CG	LEU	A	325	25.592	20.785	46.907	1.00	28.88
1230	CD1	LEU	A	325	24.072	20.701	46.960	1.00	29.23
1231	CD2	LEU	A	325	26.163	21.105	48.276	1.00	28.09
1232	N	HIS	A	326	28.219	17.033	46.821	1.00	26.59
1233	CA	HIS	A	326	28.430	15.618	46.546	1.00	25.79
1234	C	HIS	A	326	27.003	15.162	46.264	1.00	25.33
1235	O	HIS	A	326	26.052	15.819	46.695	1.00	23.44
1236	CB	HIS	A	326	28.935	14.882	47.788	1.00	27.17
1237	CG	HIS	A	326	30.303	15.294	48.231	1.00	27.36
1238	ND1	HIS	A	326	30.942	14.704	49.301	1.00	28.09
1239	CD2	HIS	A	326	31.159	16.222	47.744	1.00	28.85
1240	CE1	HIS	A	326	32.135	15.251	49.453	1.00	28.02
1241	NE2	HIS	A	326	32.292	16.174	48.521	1.00	29.20
1242	N	GLU	A	327	26.839	14.054	45.554	1.00	24.49
1243	CA	GLU	A	327	25.497	13.569	45.267	1.00	24.94
1244	C	GLU	A	327	24.768	13.297	46.583	1.00	24.29
1245	O	GLU	A	327	23.553	13.498	46.686	1.00	24.42
1246	CB	GLU	A	327	25.557	12.302	44.409	1.00	27.30
1247	CG	GLU	A	327	24.185	11.755	44.032	1.00	29.69
1248	CD	GLU	A	327	24.247	10.740	42.903	1.00	32.63
1249	OE1	GLU	A	327	25.021	9.771	43.015	1.00	31.56
1250	OE2	GLU	A	327	23.519	10.915	41.903	1.00	32.79
1251	N	GLU	A	328	25.516	12.858	47.595	1.00	22.79
1252	CA	GLU	A	328	24.942	12.576	48.911	1.00	23.11
1253	C	GLU	A	328	24.280	13.822	49.500	1.00	23.46
1254	O	GLU	A	328	23.199	13.750	50.086	1.00	23.51
1255	CB	GLU	A	328	26.025	12.083	49.877	1.00	24.71
1256	CG	GLU	A	328	26.540	10.666	49.607	1.00	25.97
1257	CD	GLU	A	328	27.584	10.591	48.506	1.00	28.78
1258	OE1	GLU	A	328	28.201	9.512	48.356	1.00	29.81
1259	OE2	GLU	A	328	27.793	11.591	47.789	1.00	27.61
1260	N	GLU	A	329	24.939	14.965	49.349	1.00	22.58
1261	CA	GLU	A	329	24.406	16.221	49.861	1.00	23.12
1262	C	GLU	A	329	23.212	16.678	49.026	1.00	22.73
1263	O	GLU	A	329	22.236	17.203	49.558	1.00	22.26
1264	CB	GLU	A	329	25.511	17.281	49.856	1.00	23.27
1265	CG	GLU	A	329	26.608	16.943	50.859	1.00	24.71
1266	CD	GLU	A	329	27.940	17.599	50.554	1.00	26.09
1267	OE1	GLU	A	329	28.825	17.532	51.429	1.00	27.57
1268	OE2	GLU	A	329	28.113	18.160	49.454	1.00	26.68
1269	N	HIS	A	330	23.291	16.450	47.721	1.00	22.72
1270	CA	HIS	A	330	22.225	16.836	46.803	1.00	22.97
1271	C	HIS	A	330	20.908	16.139	47.150	1.00	23.43
1272	O	HIS	A	330	19.863	16.790	47.257	1.00	22.10
1273	CB	HIS	A	330	22.638	16.494	45.364	1.00	24.13
1279	N	VAL	A	331	20.955	14.823	47.334	1.00	22.22
1280	CA	VAL	A	331	19.739	14.072	47.642	1.00	23.00
1281	C	VAL	A	331	19.185	14.382	49.024	1.00	22.12

1282	O	VAL	A	331	17.968	14.393	49.218	1.00	21.17
1283	CB	VAL	A	331	19.952	12.544	47.490	1.00	22.74
1284	CG1	VAL	A	331	20.363	12.233	46.053	1.00	25.60
1285	CG2	VAL	A	331	21.008	12.045	48.466	1.00	25.97
1286	N	LEU	A	332	20.067	14.634	49.986	1.00	21.61
1287	CA	LEU	A	332	19.611	14.967	51.327	1.00	21.81
1288	C	LEU	A	332	18.884	16.311	51.301	1.00	21.82
1289	O	LEU	A	332	17.874	16.489	51.976	1.00	22.23
1290	CB	LEU	A	332	20.796	15.020	52.303	1.00	22.40
1291	CG	LEU	A	332	21.262	13.656	52.824	1.00	22.71
1292	CD1	LEU	A	332	22.617	13.777	53.516	1.00	23.21
1293	CD2	LEU	A	332	20.214	13.112	53.776	1.00	23.85
1294	N	LEU	A	333	19.389	17.253	50.508	1.00	21.45
1295	CA	LEU	A	333	18.763	18.569	50.420	1.00	22.43
1296	C	LEU	A	333	17.363	18.478	49.808	1.00	21.61
1297	O	LEU	A	333	16.440	19.157	50.259	1.00	21.39
1298	CB	LEU	A	333	19.637	19.521	49.599	1.00	23.63
1302	N	MET	A	334	17.198	17.654	48.776	1.00	21.27
1303	CA	MET	A	334	15.878	17.513	48.163	1.00	20.93
1304	C	MET	A	334	14.928	16.881	49.171	1.00	21.48
1305	O	MET	A	334	13.769	17.263	49.256	1.00	21.52
1306	CB	MET	A	334	15.939	16.648	46.896	1.00	21.53
1310	N	ALA	A	335	15.427	15.922	49.950	1.00	21.64
1311	CA	ALA	A	335	14.596	15.255	50.949	1.00	21.82
1312	C	ALA	A	335	14.167	16.231	52.045	1.00	22.81
1313	O	ALA	A	335	13.002	16.248	52.455	1.00	22.95
1314	CB	ALA	A	335	15.355	14.070	51.564	1.00	22.44
1315	N	ILE	A	336	15.111	17.041	52.517	1.00	21.90
1316	CA	ILE	A	336	14.827	18.022	53.560	1.00	22.98
1317	C	ILE	A	336	13.822	19.050	53.038	1.00	23.92
1318	O	ILE	A	336	12.949	19.496	53.772	1.00	23.55
1319	CB	ILE	A	336	16.129	18.730	54.020	1.00	23.77
1320	CG1	ILE	A	336	17.021	17.724	54.753	1.00	24.24
1321	CG2	ILE	A	336	15.803	19.914	54.936	1.00	25.02
1322	CD1	ILE	A	336	18.445	18.188	54.950	1.00	27.51
1323	N	CYS	A	337	13.942	19.411	51.765	1.00	22.82
1324	CA	CYS	A	337	13.020	20.365	51.166	1.00	23.92
1325	C	CYS	A	337	11.582	19.846	51.235	1.00	24.00
1326	O	CYS	A	337	10.665	20.577	51.605	1.00	25.45
1327	CB	CYS	A	337	13.410	20.622	49.705	1.00	22.95
1329	N	ILE	A	338	11.393	18.578	50.886	1.00	23.65
1330	CA	ILE	A	338	10.070	17.957	50.890	1.00	23.56
1331	C	ILE	A	338	9.457	17.814	52.284	1.00	26.02
1332	O	ILE	A	338	8.288	18.153	52.501	1.00	25.70
1333	CB	ILE	A	338	10.126	16.560	50.231	1.00	23.28
1337	N	VAL	A	339	10.242	17.305	53.225	1.00	26.29
1338	CA	VAL	A	339	9.754	17.106	54.584	1.00	29.21
1339	C	VAL	A	339	9.971	18.359	55.430	1.00	29.45
1340	O	VAL	A	339	10.807	18.378	56.333	1.00	30.57
1341	CB	VAL	A	339	10.461	15.901	55.241	1.00	30.37
1344	N	SER	A	340	9.213	19.407	55.122	1.00	30.04
1345	CA	SER	A	340	9.309	20.676	55.842	1.00	30.94
1346	C	SER	A	340	8.061	20.868	56.701	1.00	31.69
1347	O	SER	A	340	6.940	20.841	56.195	1.00	31.64
1348	CB	SER	A	340	9.438	21.838	54.853	1.00	32.39

1349	OG	SER	A	340	10.664	21.773	54.142	1.00	35.78
1350	N	PRO	A	341	8.243	21.075	58.013	1.00	32.38
1351	CA	PRO	A	341	7.107	21.263	58.919	1.00	33.82
1352	C	PRO	A	341	6.344	22.579	58.774	1.00	35.40
1353	O	PRO	A	341	5.204	22.688	59.232	1.00	36.23
1354	CB	PRO	A	341	7.745	21.111	60.298	1.00	33.68
1355	CG	PRO	A	341	9.110	21.675	60.094	1.00	33.78
1356	CD	PRO	A	341	9.517	21.088	58.754	1.00	32.78
1357	N	ASP	A	342	6.954	23.570	58.131	1.00	36.36
1358	CA	ASP	A	342	6.301	24.866	57.981	1.00	37.82
1359	C	ASP	A	342	5.580	25.094	56.657	1.00	38.30
1360	O	ASP	A	342	5.655	26.181	56.084	1.00	39.93
1361	CB	ASP	A	342	7.304	26.001	58.213	1.00	39.62
1362	CG	ASP	A	342	8.441	25.987	57.218	1.00	41.16
1363	OD1	ASP	A	342	9.185	26.989	57.152	1.00	43.27
1364	OD2	ASP	A	342	8.597	24.974	56.505	1.00	42.71
1365	N	ARG	A	343	4.887	24.072	56.170	1.00	37.31
1366	CA	ARG	A	343	4.123	24.195	54.933	1.00	37.10
1367	C	ARG	A	343	2.683	24.409	55.375	1.00	37.98
1368	O	ARG	A	343	2.198	23.723	56.273	1.00	38.15
1369	CB	ARG	A	343	4.207	22.911	54.103	1.00	35.71
1370	CG	ARG	A	343	5.595	22.543	53.612	1.00	32.83
1371	CD	ARG	A	343	6.123	23.513	52.565	1.00	31.70
1372	NE	ARG	A	343	7.282	22.947	51.879	1.00	29.77
1373	CZ	ARG	A	343	8.062	23.612	51.032	1.00	28.95
1374	NH1	ARG	A	343	7.818	24.886	50.751	1.00	27.66
1375	NH2	ARG	A	343	9.097	23.002	50.472	1.00	28.12
1376	N	PRO	A	344	1.979	25.367	54.761	1.00	39.05
1377	CA	PRO	A	344	0.592	25.579	55.180	1.00	39.38
1378	C	PRO	A	344	-0.279	24.351	54.924	1.00	38.87
1379	O	PRO	A	344	-0.205	23.742	53.858	1.00	39.29
1380	CB	PRO	A	344	0.167	26.786	54.348	1.00	39.90
1381	CG	PRO	A	344	0.974	26.624	53.092	1.00	40.56
1382	CD	PRO	A	344	2.332	26.243	53.630	1.00	39.66
1383	N	GLY	A	345	-1.085	23.978	55.912	1.00	38.50
1384	CA	GLY	A	345	-1.965	22.836	55.746	1.00	37.96
1385	C	GLY	A	345	-1.567	21.544	56.437	1.00	37.94
1386	O	GLY	A	345	-2.386	20.630	56.537	1.00	36.75
1387	N	VAL	A	346	-0.328	21.452	56.914	1.00	37.90
1388	CA	VAL	A	346	0.125	20.234	57.585	1.00	38.39
1389	C	VAL	A	346	-0.584	20.046	58.922	1.00	39.30
1390	O	VAL	A	346	-0.832	21.012	59.643	1.00	39.35
1391	CB	VAL	A	346	1.654	20.249	57.827	1.00	38.81
1392	CG1	VAL	A	346	2.383	20.409	56.503	1.00	37.92
1393	CG2	VAL	A	346	2.030	21.366	58.784	1.00	38.02
1394	N	GLN	A	347	-0.905	18.796	59.247	1.00	39.79
1395	CA	GLN	A	347	-1.597	18.481	60.492	1.00	40.19
1396	C	GLN	A	347	-0.631	18.135	61.612	1.00	38.86
1397	O	GLN	A	347	-0.657	18.758	62.673	1.00	39.58
1398	CB	GLN	A	347	-2.564	17.314	60.280	1.00	43.00
1403	N	ASP	A	348	0.223	17.144	61.380	1.00	36.34
1404	CA	ASP	A	348	1.181	16.730	62.398	1.00	35.34
1405	C	ASP	A	348	2.568	17.313	62.152	1.00	33.83
1406	O	ASP	A	348	3.474	16.622	61.679	1.00	33.85
1407	CB	ASP	A	348	1.257	15.203	62.458	1.00	34.84

1411	N	ALA	A	349	2.727	18.587	62.492	1.00	32.41
1412	CA	ALA	A	349	3.991	19.286	62.307	1.00	32.51
1413	C	ALA	A	349	5.122	18.665	63.121	1.00	32.68
1414	O	ALA	A	349	6.263	18.602	62.662	1.00	32.47
1415	CB	ALA	A	349	3.829	20.753	62.677	1.00	32.86
1416	N	ALA	A	350	4.804	18.206	64.328	1.00	31.95
1417	CA	ALA	A	350	5.809	17.602	65.200	1.00	31.15
1418	C	ALA	A	350	6.458	16.367	64.578	1.00	30.76
1419	O	ALA	A	350	7.676	16.190	64.655	1.00	30.22
1420	CB	ALA	A	350	5.180	17.240	66.547	1.00	32.37
1421	N	LEU	A	351	5.643	15.510	63.972	1.00	30.64
1422	CA	LEU	A	351	6.150	14.298	63.340	1.00	30.92
1423	C	LEU	A	351	7.032	14.690	62.156	1.00	30.72
1424	O	LEU	A	351	8.137	14.181	61.995	1.00	30.35
1425	CB	LEU	A	351	4.989	13.428	62.848	1.00	32.92
1429	N	ILE	A	352	6.531	15.597	61.325	1.00	30.54
1430	CA	ILE	A	352	7.282	16.056	60.158	1.00	29.35
1431	C	ILE	A	352	8.628	16.646	60.580	1.00	29.73
1432	O	ILE	A	352	9.658	16.371	59.959	1.00	30.37
1433	CB	ILE	A	352	6.465	17.107	59.362	1.00	29.44
1437	N	GLU	A	353	8.626	17.445	61.644	1.00	30.02
1438	CA	GLU	A	353	9.857	18.058	62.130	1.00	30.56
1439	C	GLU	A	353	10.845	17.000	62.613	1.00	29.99
1440	O	GLU	A	353	12.050	17.147	62.438	1.00	29.97
1441	CB	GLU	A	353	9.565	19.048	63.266	1.00	32.59
1446	N	ALA	A	354	10.334	15.935	63.223	1.00	29.90
1447	CA	ALA	A	354	11.191	14.861	63.716	1.00	29.96
1448	C	ALA	A	354	11.871	14.191	62.531	1.00	29.90
1449	O	ALA	A	354	13.064	13.904	62.570	1.00	31.40
1450	CB	ALA	A	354	10.367	13.843	64.491	1.00	30.18
1451	N	ILE	A	355	11.100	13.940	61.478	1.00	29.82
1452	CA	ILE	A	355	11.638	13.314	60.274	1.00	28.57
1453	C	ILE	A	355	12.687	14.220	59.628	1.00	28.03
1454	O	ILE	A	355	13.754	13.756	59.234	1.00	27.72
1455	CB	ILE	A	355	10.514	13.022	59.259	1.00	29.38
1459	N	GLN	A	356	12.398	15.515	59.534	1.00	27.45
1460	CA	GLN	A	356	13.345	16.444	58.925	1.00	28.48
1461	C	GLN	A	356	14.621	16.566	59.754	1.00	29.24
1462	O	GLN	A	356	15.719	16.622	59.202	1.00	27.50
1463	CB	GLN	A	356	12.718	17.833	58.739	1.00	28.93
1464	CG	GLN	A	356	13.536	18.753	57.823	1.00	29.68
1465	CD	GLN	A	356	13.064	20.198	57.844	1.00	31.36
1466	OE1	GLN	A	356	12.996	20.823	58.903	1.00	31.51
1467	NE2	GLN	A	356	12.747	20.742	56.667	1.00	30.04
1468	N	ASP	A	357	14.482	16.613	61.078	1.00	29.52
1469	CA	ASP	A	357	15.656	16.724	61.945	1.00	30.44
1470	C	ASP	A	357	16.610	15.550	61.739	1.00	29.70
1471	O	ASP	A	357	17.827	15.729	61.727	1.00	30.37
1472	CB	ASP	A	357	15.244	16.791	63.423	1.00	32.83
1476	N	ARG	A	358	16.059	14.351	61.577	1.00	30.08
1477	CA	ARG	A	358	16.887	13.167	61.368	1.00	30.11
1478	C	ARG	A	358	17.689	13.309	60.073	1.00	29.92
1479	O	ARG	A	358	18.842	12.880	59.996	1.00	29.08
1480	CB	ARG	A	358	16.014	11.906	61.323	1.00	31.17
1487	N	LEU	A	359	17.074	13.919	59.061	1.00	28.41

1488	CA	LEU	A	359	17.735	14.128	57.776	1.00	27.76
1489	C	LEU	A	359	18.757	15.255	57.890	1.00	27.89
1490	O	LEU	A	359	19.853	15.171	57.338	1.00	27.13
1491	CB	LEU	A	359	16.704	14.482	56.697	1.00	27.85
1495	N	SER	A	360	18.393	16.312	58.610	1.00	28.62
1496	CA	SER	A	360	19.288	17.448	58.790	1.00	30.04
1497	C	SER	A	360	20.540	17.046	59.561	1.00	30.70
1498	O	SER	A	360	21.647	17.454	59.212	1.00	31.03
1499	CB	SER	A	360	18.573	18.578	59.534	1.00	32.12
1501	N	ASN	A	361	20.367	16.251	60.613	1.00	31.34
1502	CA	ASN	A	361	21.513	15.816	61.405	1.00	31.58
1503	C	ASN	A	361	22.417	14.921	60.570	1.00	30.77
1504	O	ASN	A	361	23.637	14.935	60.728	1.00	31.06
1505	CB	ASN	A	361	21.055	15.083	62.667	1.00	34.28
1509	N	THR	A	362	21.815	14.146	59.674	1.00	29.26
1510	CA	THR	A	362	22.583	13.270	58.800	1.00	28.14
1511	C	THR	A	362	23.419	14.135	57.863	1.00	27.56
1512	O	THR	A	362	24.607	13.879	57.654	1.00	27.15
1513	CB	THR	A	362	21.654	12.371	57.956	1.00	28.47
1514	OG1	THR	A	362	20.923	11.495	58.823	1.00	28.00
1515	CG2	THR	A	362	22.461	11.548	56.955	1.00	27.60
1516	N	LEU	A	363	22.795	15.167	57.301	1.00	26.97
1517	CA	LEU	A	363	23.493	16.064	56.388	1.00	27.40
1518	C	LEU	A	363	24.623	16.798	57.100	1.00	28.23
1519	O	LEU	A	363	25.736	16.884	56.588	1.00	27.96
1520	CB	LEU	A	363	22.519	17.089	55.782	1.00	26.59
1521	CG	LEU	A	363	23.153	18.156	54.882	1.00	26.54
1522	CD1	LEU	A	363	23.829	17.495	53.687	1.00	26.43
1523	CD2	LEU	A	363	22.090	19.142	54.417	1.00	26.28
1524	N	GLN	A	364	24.340	17.325	58.286	1.00	29.48
1525	CA	GLN	A	364	25.360	18.054	59.029	1.00	31.77
1526	C	GLN	A	364	26.530	17.140	59.399	1.00	30.91
1527	O	GLN	A	364	27.691	17.539	59.307	1.00	30.91
1528	CB	GLN	A	364	24.747	18.681	60.283	1.00	33.97
1533	N	THR	A	365	26.224	15.910	59.799	1.00	30.38
1534	CA	THR	A	365	27.263	14.956	60.176	1.00	30.54
1535	C	THR	A	365	28.099	14.561	58.965	1.00	29.67
1536	O	THR	A	365	29.319	14.454	59.054	1.00	30.84
1537	CB	THR	A	365	26.658	13.687	60.802	1.00	30.66
1540	N	TYR	A	366	27.437	14.348	57.832	1.00	29.11
1541	CA	TYR	A	366	28.131	13.976	56.606	1.00	28.71
1542	C	TYR	A	366	29.123	15.051	56.167	1.00	28.75
1543	O	TYR	A	366	30.261	14.746	55.826	1.00	29.49
1544	CB	TYR	A	366	27.122	13.708	55.476	1.00	27.69
1545	CG	TYR	A	366	27.779	13.396	54.148	1.00	27.26
1546	CD1	TYR	A	366	28.234	14.421	53.313	1.00	27.40
1547	CD2	TYR	A	366	28.017	12.079	53.759	1.00	27.61
1548	CE1	TYR	A	366	28.912	14.144	52.130	1.00	28.23
1549	CE2	TYR	A	366	28.697	11.790	52.578	1.00	28.00
1550	CZ	TYR	A	366	29.143	12.825	51.770	1.00	28.28
1551	OH	TYR	A	366	29.838	12.546	50.615	1.00	28.34
1552	N	ILE	A	367	28.692	16.310	56.174	1.00	30.76
1553	CA	ILE	A	367	29.559	17.412	55.762	1.00	32.74
1554	C	ILE	A	367	30.823	17.533	56.614	1.00	35.15
1555	O	ILE	A	367	31.924	17.688	56.086	1.00	35.56

1556	CB	ILE	A	367	28.805	18.763	55.807	1.00	32.16
1557	CG1	ILE	A	367	27.685	18.764	54.763	1.00	32.04
1558	CG2	ILE	A	367	29.769	19.915	55.535	1.00	32.16
1559	CD1	ILE	A	367	26.790	19.977	54.829	1.00	32.66
1560	N	ARG	A	368	30.660	17.465	57.930	1.00	38.07
1561	CA	ARG	A	368	31.794	17.582	58.842	1.00	41.56
1562	C	ARG	A	368	32.749	16.406	58.711	1.00	43.08
1563	O	ARG	A	368	33.963	16.558	58.845	1.00	43.32
1564	CB	ARG	A	368	31.309	17.652	60.289	1.00	43.01
1571	N	CYS	A	369	32.187	15.234	58.440	1.00	44.44
1572	CA	CYS	A	369	32.964	14.011	58.333	1.00	46.85
1573	C	CYS	A	369	33.501	13.644	56.949	1.00	46.74
1574	O	CYS	A	369	34.641	13.198	56.828	1.00	46.26
1575	CB	CYS	A	369	32.128	12.848	58.881	1.00	48.76
1576	SG	CYS	A	369	32.925	11.238	58.816	1.00	56.08
1577	N	ARG	A	370	32.700	13.841	55.905	1.00	47.07
1578	CA	ARG	A	370	33.123	13.457	54.558	1.00	47.21
1579	C	ARG	A	370	33.451	14.563	53.559	1.00	47.24
1580	O	ARG	A	370	34.058	14.292	52.520	1.00	46.93
1581	CB	ARG	A	370	32.068	12.533	53.940	1.00	47.76
1588	N	HIS	A	371	33.059	15.799	53.845	1.00	46.83
1589	CA	HIS	A	371	33.340	16.879	52.908	1.00	47.01
1590	C	HIS	A	371	34.670	17.554	53.217	1.00	47.99
1591	O	HIS	A	371	34.809	18.227	54.237	1.00	46.99
1592	CB	HIS	A	371	32.225	17.924	52.927	1.00	45.11
1598	N	PRO	A	372	35.665	17.383	52.331	1.00	49.67
1599	CA	PRO	A	372	36.998	17.972	52.497	1.00	51.25
1600	C	PRO	A	372	37.027	19.480	52.256	1.00	52.68
1601	O	PRO	A	372	36.223	20.013	51.489	1.00	52.39
1602	CB	PRO	A	372	37.833	17.208	51.476	1.00	51.31
1603	CG	PRO	A	372	36.853	16.991	50.366	1.00	51.05
1604	CD	PRO	A	372	35.613	16.549	51.115	1.00	50.21
1605	N	PRO	A	373	37.961	20.187	52.914	1.00	54.03
1606	CA	PRO	A	373	38.107	21.640	52.777	1.00	55.28
1607	C	PRO	A	373	38.693	22.028	51.420	1.00	56.14
1608	O	PRO	A	373	39.284	21.197	50.731	1.00	56.44
1609	CB	PRO	A	373	39.036	21.998	53.932	1.00	55.37
1612	N	PRO	A	374	38.535	23.299	51.017	1.00	56.81
1613	CA	PRO	A	374	37.848	24.368	51.750	1.00	57.49
1614	C	PRO	A	374	36.324	24.301	51.617	1.00	57.83
1615	O	PRO	A	374	35.642	24.299	52.664	1.00	58.58
1616	CB	PRO	A	374	38.431	25.631	51.127	1.00	57.24
1617	CG	PRO	A	374	38.601	25.226	49.698	1.00	57.37
1618	CD	PRO	A	374	39.194	23.834	49.811	1.00	57.05
1619	N	LEU	A	378	30.279	26.156	57.018	1.00	49.88
1620	CA	LEU	A	378	29.679	27.221	56.220	1.00	45.66
1621	C	LEU	A	378	28.825	26.586	55.127	1.00	41.60
1622	O	LEU	A	378	27.802	27.138	54.723	1.00	38.14
1623	CB	LEU	A	378	30.769	28.092	55.590	1.00	53.84
1627	N	LEU	A	379	29.370	25.200	54.660	1.00	35.28
1628	CA	LEU	A	379	28.529	24.615	53.626	1.00	33.21
1629	C	LEU	A	379	27.095	24.355	54.080	1.00	32.24
1630	O	LEU	A	379	26.157	24.594	53.325	1.00	31.09
1631	CB	LEU	A	379	29.151	23.309	53.121	1.00	33.21
1632	CG	LEU	A	379	28.379	22.603	52.003	1.00	31.83

1633	CD1	LEU	A	379	28.301	23.508	50.783	1.00	33.04
1634	CD2	LEU	A	379	29.066	21.292	51.651	1.00	32.12
1635	N	TYR	A	380	26.917	23.869	55.304	1.00	31.72
1636	CA	TYR	A	380	25.572	23.588	55.792	1.00	32.34
1637	C	TYR	A	380	24.717	24.852	55.780	1.00	32.63
1638	O	TYR	A	380	23.562	24.833	55.339	1.00	31.56
1639	CB	TYR	A	380	25.611	23.008	57.208	1.00	33.03
1647	N	ALA	A	381	25.288	25.950	56.266	1.00	32.03
1648	CA	ALA	A	381	24.578	27.223	56.304	1.00	32.04
1649	C	ALA	A	381	24.190	27.683	54.902	1.00	31.59
1650	O	ALA	A	381	23.084	28.187	54.693	1.00	32.20
1651	CB	ALA	A	381	25.443	28.287	56.981	1.00	32.84
1652	N	LYS	A	382	25.101	27.515	53.948	1.00	30.09
1653	CA	LYS	A	382	24.849	27.916	52.570	1.00	30.96
1654	C	LYS	A	382	23.739	27.083	51.943	1.00	30.08
1655	O	LYS	A	382	22.989	27.575	51.101	1.00	30.60
1656	CB	LYS	A	382	26.121	27.781	51.731	1.00	31.98
1657	CG	LYS	A	382	27.223	28.757	52.109	1.00	34.76
1658	CD	LYS	A	382	28.458	28.545	51.254	1.00	38.05
1659	CE	LYS	A	382	29.559	29.526	51.615	1.00	39.31
1660	NZ	LYS	A	382	30.806	29.245	50.845	1.00	41.47
1661	N	MET	A	383	23.648	25.819	52.345	1.00	29.65
1662	CA	MET	A	383	22.621	24.923	51.821	1.00	29.41
1663	C	MET	A	383	21.253	25.286	52.389	1.00	29.82
1664	O	MET	A	383	20.250	25.271	51.677	1.00	29.12
1665	CB	MET	A	383	22.958	23.468	52.165	1.00	28.17
1669	N	ILE	A	384	21.215	25.612	53.676	1.00	30.76
1670	CA	ILE	A	384	19.960	25.983	54.319	1.00	32.84
1671	C	ILE	A	384	19.422	27.271	53.701	1.00	32.96
1672	O	ILE	A	384	18.208	27.458	53.594	1.00	32.83
1673	CB	ILE	A	384	20.149	26.186	55.842	1.00	34.53
1677	N	GLN	A	385	20.328	28.153	53.287	1.00	32.82
1678	CA	GLN	A	385	19.931	29.412	52.669	1.00	33.03
1679	C	GLN	A	385	19.288	29.174	51.303	1.00	32.26
1680	O	GLN	A	385	18.382	29.905	50.901	1.00	30.38
1681	CB	GLN	A	385	21.136	30.342	52.515	1.00	35.19
1686	N	LYS	A	386	19.756	28.152	50.591	1.00	30.96
1687	CA	LYS	A	386	19.197	27.840	49.282	1.00	30.88
1688	C	LYS	A	386	17.748	27.415	49.447	1.00	29.98
1689	O	LYS	A	386	16.927	27.635	48.558	1.00	29.72
1690	CB	LYS	A	386	19.985	26.719	48.601	1.00	32.29
1695	N	LEU	A	387	17.433	26.804	50.583	1.00	29.25
1696	CA	LEU	A	387	16.064	26.373	50.833	1.00	29.50
1697	C	LEU	A	387	15.172	27.604	50.982	1.00	29.30
1698	O	LEU	A	387	14.014	27.594	50.572	1.00	27.98
1699	CB	LEU	A	387	15.988	25.503	52.091	1.00	30.49
1700	CG	LEU	A	387	16.625	24.111	51.980	1.00	31.01
1701	CD1	LEU	A	387	16.443	23.363	53.289	1.00	33.01
1702	CD2	LEU	A	387	15.985	23.339	50.839	1.00	31.27
1703	N	ALA	A	388	15.714	28.667	51.566	1.00	28.72
1704	CA	ALA	A	388	14.952	29.903	51.735	1.00	29.54
1705	C	ALA	A	388	14.757	30.557	50.367	1.00	29.67
1706	O	ALA	A	388	13.696	31.121	50.082	1.00	29.63
1707	CB	ALA	A	388	15.687	30.856	52.679	1.00	30.01
1708	N	ASP	A	389	15.786	30.479	49.524	1.00	29.62

1709	CA	ASP	A	389	15.730	31.044	48.175	1.00	30.08
1710	C	ASP	A	389	14.625	30.360	47.378	1.00	29.78
1711	O	ASP	A	389	13.917	31.000	46.598	1.00	29.18
1712	CB	ASP	A	389	17.059	30.833	47.445	1.00	31.12
1716	N	LEU	A	390	14.496	29.052	47.572	1.00	27.64
1717	CA	LEU	A	390	13.482	28.262	46.881	1.00	28.50
1718	C	LEU	A	390	12.067	28.730	47.223	1.00	27.80
1719	O	LEU	A	390	11.187	28.740	46.360	1.00	26.98
1720	CB	LEU	A	390	13.646	26.786	47.252	1.00	28.87
1721	CG	LEU	A	390	14.130	25.761	46.221	1.00	31.08
1722	CD1	LEU	A	390	14.754	26.424	45.017	1.00	31.00
1723	CD2	LEU	A	390	15.101	24.810	46.902	1.00	31.75
1724	N	ARG	A	391	11.849	29.109	48.481	1.00	27.87
1725	CA	ARG	A	391	10.535	29.574	48.917	1.00	28.52
1726	C	ARG	A	391	10.132	30.808	48.125	1.00	28.78
1727	O	ARG	A	391	8.968	30.961	47.757	1.00	28.82
1728	CB	ARG	A	391	10.536	29.919	50.415	1.00	30.35
1729	CG	ARG	A	391	10.795	28.744	51.354	1.00	32.51
1730	CD	ARG	A	391	9.743	27.658	51.208	1.00	34.99
1731	NE	ARG	A	391	9.952	26.552	52.141	1.00	37.12
1732	CZ	ARG	A	391	9.395	26.460	53.346	1.00	37.91
1733	NH1	ARG	A	391	8.580	27.411	53.783	1.00	38.94
1734	NH2	ARG	A	391	9.646	25.408	54.115	1.00	37.42
1735	N	SER	A	392	11.094	31.690	47.865	1.00	28.76
1736	CA	SER	A	392	10.811	32.908	47.114	1.00	29.36
1737	C	SER	A	392	10.483	32.588	45.664	1.00	28.02
1738	O	SER	A	392	9.577	33.178	45.082	1.00	28.38
1739	CB	SER	A	392	11.997	33.866	47.185	1.00	31.21
1740	OG	SER	A	392	12.192	34.305	48.518	1.00	37.19
1741	N	LEU	A	393	11.219	31.648	45.081	1.00	26.23
1742	CA	LEU	A	393	10.972	31.253	43.700	1.00	26.10
1743	C	LEU	A	393	9.614	30.567	43.586	1.00	25.57
1744	O	LEU	A	393	8.919	30.705	42.576	1.00	26.87
1745	CB	LEU	A	393	12.081	30.309	43.216	1.00	26.02
1749	N	ASN	A	394	9.242	29.825	44.625	1.00	24.50
1750	CA	ASN	A	394	7.964	29.122	44.656	1.00	26.07
1751	C	ASN	A	394	6.855	30.167	44.570	1.00	27.28
1752	O	ASN	A	394	5.929	30.055	43.764	1.00	26.29
1753	CB	ASN	A	394	7.827	28.347	45.967	1.00	26.75
1754	CG	ASN	A	394	6.646	27.397	45.968	1.00	28.26
1755	OD1	ASN	A	394	5.660	27.604	45.263	1.00	28.24
1756	ND2	ASN	A	394	6.736	26.352	46.779	1.00	28.79
1757	N	GLU	A	395	6.966	31.188	45.413	1.00	28.62
1758	CA	GLU	A	395	5.986	32.266	45.464	1.00	30.55
1759	C	GLU	A	395	5.815	32.976	44.130	1.00	29.66
1760	O	GLU	A	395	4.691	33.213	43.684	1.00	29.50
1761	CB	GLU	A	395	6.385	33.280	46.536	1.00	33.44
1766	N	GLU	A	396	6.929	33.324	43.496	1.00	29.08
1767	CA	GLU	A	396	6.871	34.013	42.217	1.00	28.78
1768	C	GLU	A	396	6.280	33.102	41.148	1.00	28.20
1769	O	GLU	A	396	5.486	33.545	40.317	1.00	27.96
1770	CB	GLU	A	396	8.265	34.490	41.791	1.00	30.45
1775	N	HIS	A	397	6.651	31.826	41.162	1.00	26.94
1776	CA	HIS	A	397	6.104	30.919	40.162	1.00	27.05
1777	C	HIS	A	397	4.583	30.835	40.295	1.00	27.50

1778	O	HIS	A	397	3.866	30.834	39.294	1.00	27.05
1779	CB	HIS	A	397	6.718	29.519	40.282	1.00	26.64
1780	CG	HIS	A	397	6.058	28.507	39.400	1.00	26.04
1781	ND1	HIS	A	397	4.999	27.731	39.822	1.00	27.22
1782	CD2	HIS	A	397	6.227	28.228	38.086	1.00	26.36
1783	CE1	HIS	A	397	4.542	27.024	38.805	1.00	26.59
1784	NE2	HIS	A	397	5.268	27.308	37.740	1.00	26.30
1785	N	SER	A	398	4.094	30.785	41.529	1.00	28.04
1786	CA	SER	A	398	2.657	30.696	41.775	1.00	29.62
1787	C	SER	A	398	1.921	31.901	41.195	1.00	29.87
1788	O	SER	A	398	0.862	31.761	40.579	1.00	28.32
1789	CB	SER	A	398	2.389	30.604	43.279	1.00	31.44
1791	N	LYS	A	399	2.485	33.085	41.397	1.00	30.06
1792	CA	LYS	A	399	1.882	34.313	40.885	1.00	31.61
1793	C	LYS	A	399	1.807	34.283	39.363	1.00	30.91
1794	O	LYS	A	399	0.790	34.651	38.771	1.00	30.21
1795	CB	LYS	A	399	2.698	35.527	41.336	1.00	33.79
1821	N	ARG	A	402	-0.742	31.790	37.948	1.00	29.27
1822	CA	ARG	A	402	-2.083	32.360	38.021	1.00	32.16
1823	C	ARG	A	402	-2.386	33.173	36.769	1.00	32.00
1824	O	ARG	A	402	-3.434	32.998	36.150	1.00	31.31
1825	CB	ARG	A	402	-2.220	33.251	39.256	1.00	36.10
1826	CG	ARG	A	402	-3.587	33.906	39.391	1.00	41.28
1827	CD	ARG	A	402	-3.710	34.730	40.669	1.00	45.66
1828	NE	ARG	A	402	-3.552	33.916	41.873	1.00	49.74
1829	CZ	ARG	A	402	-2.382	33.582	42.410	1.00	51.64
1830	NH1	ARG	A	402	-1.249	33.996	41.856	1.00	52.67
1831	NH2	ARG	A	402	-2.343	32.825	43.499	1.00	52.69
1832	N	CYS	A	403	-1.471	34.066	36.402	1.00	32.10
1833	CA	CYS	A	403	-1.645	34.895	35.210	1.00	33.04
1834	C	CYS	A	403	-1.781	34.014	33.976	1.00	32.14
1835	O	CYS	A	403	-2.620	34.257	33.106	1.00	30.55
1836	CB	CYS	A	403	-0.450	35.838	35.030	1.00	35.60
1838	N	LEU	A	404	-0.950	32.980	33.911	1.00	30.59
1839	CA	LEU	A	404	-0.967	32.065	32.784	1.00	30.86
1840	C	LEU	A	404	-2.327	31.390	32.638	1.00	29.62
1841	O	LEU	A	404	-2.840	31.256	31.529	1.00	30.21
1842	CB	LEU	A	404	0.130	31.008	32.955	1.00	32.55
1843	CG	LEU	A	404	0.353	30.078	31.766	1.00	34.87
1844	CD1	LEU	A	404	0.840	30.895	30.580	1.00	36.26
1845	CD2	LEU	A	404	1.370	29.005	32.127	1.00	35.52
1846	N	SER	A	405	-2.918	30.987	33.760	1.00	29.30
1847	CA	SER	A	405	-4.212	30.309	33.749	1.00	29.83
1848	C	SER	A	405	-5.358	31.173	33.218	1.00	28.16
1849	O	SER	A	405	-6.423	30.651	32.885	1.00	28.54
1850	CB	SER	A	405	-4.563	29.802	35.153	1.00	31.58
1852	N	PHE	A	406	-5.147	32.484	33.145	1.00	25.99
1853	CA	PHE	A	406	-6.179	33.396	32.636	1.00	26.56
1854	C	PHE	A	406	-6.263	33.340	31.112	1.00	26.23
1855	O	PHE	A	406	-7.256	33.778	30.518	1.00	25.59
1856	CB	PHE	A	406	-5.868	34.842	33.042	1.00	26.14
1863	N	GLN	A	407	-5.220	32.814	30.478	1.00	25.64
1864	CA	GLN	A	407	-5.189	32.748	29.019	1.00	25.17
1865	C	GLN	A	407	-6.155	31.687	28.500	1.00	25.33
1866	O	GLN	A	407	-6.086	30.524	28.903	1.00	24.86

1867	CB	GLN	A	407	-3.765	32.448	28.527	1.00	25.99
1868	CG	GLN	A	407	-3.571	32.694	27.030	1.00	26.23
1869	CD	GLN	A	407	-3.718	34.165	26.651	1.00	26.81
1870	OE1	GLN	A	407	-4.087	34.494	25.520	1.00	28.94
1871	NE2	GLN	A	407	-3.414	35.052	27.590	1.00	21.63
1872	N	PRO	A	408	-7.083	32.079	27.608	1.00	25.83
1873	CA	PRO	A	408	-8.052	31.124	27.058	1.00	27.42
1874	C	PRO	A	408	-7.384	29.913	26.398	1.00	29.12
1875	O	PRO	A	408	-6.389	30.056	25.688	1.00	29.12
1876	CB	PRO	A	408	-8.835	31.967	26.054	1.00	25.96
1879	N	GLU	A	409	-7.941	28.731	26.646	1.00	31.66
1880	CA	GLU	A	409	-7.441	27.479	26.078	1.00	34.64
1881	C	GLU	A	409	-6.104	27.014	26.661	1.00	34.56
1882	O	GLU	A	409	-5.480	26.100	26.122	1.00	34.24
1883	CB	GLU	A	409	-7.293	27.606	24.555	1.00	37.88
1888	N	CYS	A	410	-5.671	27.628	27.759	1.00	33.81
1889	CA	CYS	A	410	-4.399	27.267	28.382	1.00	35.08
1890	C	CYS	A	410	-4.396	25.871	29.002	1.00	34.78
1891	O	CYS	A	410	-3.390	25.164	28.943	1.00	34.24
1892	CB	CYS	A	410	-4.027	28.299	29.455	1.00	36.13
1894	N	SER	A	411	-5.518	25.472	29.593	1.00	34.62
1895	CA	SER	A	411	-5.611	24.163	30.235	1.00	35.60
1896	C	SER	A	411	-5.215	23.008	29.319	1.00	35.51
1897	O	SER	A	411	-4.602	22.040	29.770	1.00	35.62
1898	CB	SER	A	411	-7.031	23.928	30.763	1.00	36.58
1900	N	MET	A	412	-5.561	23.108	28.038	1.00	35.70
1901	CA	MET	A	412	-5.244	22.053	27.079	1.00	36.25
1902	C	MET	A	412	-3.744	21.912	26.846	1.00	34.53
1903	O	MET	A	412	-3.273	20.867	26.393	1.00	34.48
1904	CB	MET	A	412	-5.936	22.324	25.741	1.00	40.58
1908	N	LYS	A	413	-2.996	22.965	27.150	1.00	31.53
1909	CA	LYS	A	413	-1.551	22.944	26.960	1.00	30.85
1910	C	LYS	A	413	-0.831	22.407	28.192	1.00	30.52
1911	O	LYS	A	413	0.386	22.236	28.187	1.00	30.68
1912	CB	LYS	A	413	-1.042	24.350	26.632	1.00	31.05
1917	N	LEU	A	414	-1.590	22.142	29.248	1.00	30.09
1918	CA	LEU	A	414	-1.014	21.620	30.484	1.00	28.96
1919	C	LEU	A	414	-1.393	20.147	30.610	1.00	28.33
1920	O	LEU	A	414	-1.654	19.489	29.604	1.00	29.10
1921	CB	LEU	A	414	-1.544	22.427	31.676	1.00	28.64
1925	N	THR	A	415	-1.401	19.624	31.833	1.00	27.70
1926	CA	THR	A	415	-1.779	18.232	32.071	1.00	26.69
1927	C	THR	A	415	-2.620	18.195	33.338	1.00	26.91
1928	O	THR	A	415	-2.548	19.104	34.157	1.00	26.39
1929	CB	THR	A	415	-0.556	17.310	32.307	1.00	26.61
1930	OG1	THR	A	415	-0.006	17.570	33.607	1.00	25.35
1931	CG2	THR	A	415	0.509	17.546	31.247	1.00	26.48
1932	N	PRO	A	416	-3.432	17.142	33.516	1.00	27.60
1933	CA	PRO	A	416	-4.269	17.037	34.717	1.00	27.16
1934	C	PRO	A	416	-3.477	17.169	36.026	1.00	27.48
1935	O	PRO	A	416	-3.930	17.813	36.975	1.00	26.90
1936	CB	PRO	A	416	-4.908	15.661	34.564	1.00	29.00
1937	CG	PRO	A	416	-5.083	15.555	33.072	1.00	28.36
1938	CD	PRO	A	416	-3.752	16.071	32.553	1.00	28.22
1939	N	LEU	A	417	-2.294	16.560	36.072	1.00	25.90

1940	CA	LEU	A	417	-1.460	16.610	37.271	1.00	25.39
1941	C	LEU	A	417	-0.961	18.031	37.545	1.00	24.67
1942	O	LEU	A	417	-0.983	18.502	38.685	1.00	24.55
1943	CB	LEU	A	417	-0.279	15.643	37.124	1.00	25.12
1944	CG	LEU	A	417	0.722	15.507	38.273	1.00	25.26
1945	CD1	LEU	A	417	0.021	15.098	39.564	1.00	24.40
1946	CD2	LEU	A	417	1.766	14.470	37.882	1.00	25.23
1947	N	VAL	A	418	-0.506	18.711	36.500	1.00	24.66
1948	CA	VAL	A	418	-0.027	20.080	36.640	1.00	25.57
1949	C	VAL	A	418	-1.176	20.971	37.111	1.00	26.33
1950	O	VAL	A	418	-1.001	21.814	37.991	1.00	27.09
1951	CB	VAL	A	418	0.531	20.599	35.297	1.00	25.14
1954	N	LEU	A	419	-2.354	20.769	36.530	1.00	26.33
1955	CA	LEU	A	419	-3.526	21.556	36.902	1.00	27.78
1956	C	LEU	A	419	-3.861	21.399	38.382	1.00	29.03
1957	O	LEU	A	419	-4.206	22.370	39.052	1.00	30.30
1958	CB	LEU	A	419	-4.733	21.143	36.051	1.00	28.60
1962	N	GLU	A	420	-3.738	20.184	38.904	1.00	29.76
1963	CA	GLU	A	420	-4.056	19.962	40.307	1.00	31.06
1964	C	GLU	A	420	-3.010	20.514	41.268	1.00	30.59
1965	O	GLU	A	420	-3.344	21.184	42.245	1.00	30.30
1966	CB	GLU	A	420	-4.237	18.478	40.605	1.00	32.62
1967	CG	GLU	A	420	-4.697	18.251	42.037	1.00	36.69
1968	CD	GLU	A	420	-4.267	16.919	42.598	1.00	38.47
1969	OE1	GLU	A	420	-4.631	16.624	43.756	1.00	40.46
1970	OE2	GLU	A	420	-3.561	16.171	41.891	1.00	41.56
1971	N	VAL	A	421	-1.744	20.223	40.992	1.00	31.25
1972	CA	VAL	A	421	-0.663	20.675	41.855	1.00	32.00
1973	C	VAL	A	421	-0.544	22.191	41.960	1.00	32.63
1974	O	VAL	A	421	-0.355	22.724	43.051	1.00	32.82
1975	CB	VAL	A	421	0.694	20.082	41.395	1.00	31.60
1978	N	PHE	A	422	-0.670	22.890	40.839	1.00	33.47
1979	CA	PHE	A	422	-0.541	24.342	40.857	1.00	34.95
1980	C	PHE	A	422	-1.866	25.089	40.872	1.00	35.99
1981	O	PHE	A	422	-1.907	26.284	41.159	1.00	36.37
1982	CB	PHE	A	422	0.310	24.794	39.670	1.00	34.60
1983	CG	PHE	A	422	1.679	24.182	39.656	1.00	34.76
1984	CD1	PHE	A	422	2.093	23.389	38.592	1.00	34.88
1985	CD2	PHE	A	422	2.545	24.369	40.728	1.00	35.16
1986	CE1	PHE	A	422	3.348	22.790	38.597	1.00	34.92
1987	CE2	PHE	A	422	3.801	23.774	40.743	1.00	34.70
1988	CZ	PHE	A	422	4.202	22.982	39.674	1.00	34.73
1989	N	GLY	A	423	-2.946	24.378	40.570	1.00	37.39
1990	CA	GLY	A	423	-4.261	24.993	40.564	1.00	38.89
1991	C	GLY	A	423	-4.914	24.907	41.930	1.00	39.70
1992	O	GLY	A	423	-5.857	24.099	42.083	1.00	40.52
1993	GLY	A	423						
1994	O2	VDX	425	17.029	18.071	34.819	1.00	21.73	O
1995	O3	VDX	425	4.489	26.946	35.054	1.00	24.67	O
1996	C1	VDX	425	14.139	17.953	35.755	1.00	20.80	C
1997	C2	VDX	425	14.879	16.893	34.895	1.00	21.02	C
1998	C3	VDX	425	15.992	17.534	33.962	1.00	21.41	C
1999	C4	VDX	425	15.368	18.672	33.049	1.00	21.29	C
2000	C5	VDX	425	14.622	19.724	33.864	1.00	21.00	C
2001	C6	VDX	425	14.797	21.120	33.792	1.00	20.95	C

2002	C7	VDX	425	14.174	22.286	34.514	1.00	21.23	C
2003	C8	VDX	425	13.966	23.488	34.042	1.00	21.54	C
2004	C9	VDX	425	14.354	23.927	32.544	1.00	21.77	C
2005	C10	VDX	425	13.602	19.075	34.828	1.00	20.74	C
2006	C11	VDX	425	13.088	24.490	31.671	1.00	21.66	C
2007	C12	VDX	425	12.147	25.443	32.564	1.00	22.04	C
2008	C13	VDX	425	11.753	24.897	34.070	1.00	22.01	C
2009	C14	VDX	425	13.148	24.538	34.777	1.00	21.80	C
2010	C15	VDX	425	12.661	24.266	36.350	1.00	22.22	C
2011	C16	VDX	425	11.429	25.231	36.497	1.00	22.39	C
2012	C17	VDX	425	11.276	25.934	35.106	1.00	22.31	C
2013	C18	VDX	425	10.769	23.570	33.779	1.00	21.50	C
2014	C19	VDX	425	12.291	19.455	34.852	1.00	20.77	C
2015	C20	VDX	425	9.849	26.546	34.726	1.00	22.90	C
2016	C21	VDX	425	9.804	27.956	35.482	1.00	23.65	C
2017	C22	VDX	425	8.575	25.824	35.268	1.00	23.16	C
2018	C23	VDX	425	7.331	26.060	34.405	1.00	23.73	C
2019	C24	VDX	425	6.152	25.266	34.672	1.00	24.36	C
2020	C25	VDX	425	4.775	25.776	34.336	1.00	24.75	C
2021	C26	VDX	425	4.701	26.010	32.842	1.00	25.41	C
2022	C27	VDX	425	3.668	24.730	34.723	1.00	25.39	C
2023	O1	VDX	425	13.119	17.359	36.620	1.00	20.68	O
2024	O	HOH	500	14.347	10.333	30.796	1.00	24.33	O
2025	O	HOH	501	13.828	12.782	35.922	1.00	21.46	O
2026	O	HOH	502	13.846	14.468	42.856	1.00	24.78	O
2027	O	HOH	503	19.132	15.890	40.266	1.00	21.27	O
2028	O	HOH	504	15.013	12.029	41.977	1.00	22.69	O
2029	O	HOH	505	13.766	10.118	35.125	1.00	20.29	O
2030	O	HOH	506	16.290	13.157	34.345	1.00	30.57	O
2031	O	HOH	507	5.938	22.747	23.179	1.00	24.25	O
2032	O	HOH	508	13.771	7.592	35.963	1.00	28.23	O
2033	O	HOH	509	12.348	25.386	50.763	1.00	30.93	O
2034	O	HOH	510	28.498	23.703	34.824	1.00	37.09	O
2035	O	HOH	511	26.394	10.521	64.086	1.00	30.68	O
2036	O	HOH	512	20.573	9.150	38.613	1.00	30.36	O
2037	O	HOH	513	19.724	30.629	29.203	1.00	35.40	O
2038	O	HOH	514	4.372	27.504	42.595	1.00	31.46	O
2039	O	HOH	515	2.808	13.423	33.286	1.00	30.93	O
2040	O	HOH	516	23.698	20.154	43.135	1.00	37.92	O
2041	O	HOH	517	11.325	5.901	37.588	1.00	30.12	O
2042	O	HOH	518	0.885	13.049	59.537	1.00	39.32	O
2043	O	HOH	519	20.338	11.515	62.065	1.00	36.13	O
2044	O	HOH	520	8.913	6.134	53.451	1.00	44.37	O
2045	O	HOH	521	4.924	23.321	44.129	1.00	33.51	O
2046	O	HOH	522	16.547	6.409	36.375	1.00	32.70	O
2047	O	HOH	523	8.896	35.918	45.789	1.00	45.73	O
2048	O	HOH	524	26.192	21.542	43.420	1.00	28.56	O
2049	O	HOH	525	-5.345	32.214	23.915	1.00	35.31	O
2050	O	HOH	526	9.488	15.901	22.976	1.00	29.33	O
2051	O	HOH	527	5.345	31.465	22.796	1.00	31.37	O
2052	O	HOH	528	6.982	20.227	51.589	1.00	32.20	O
2053	O	HOH	529	4.642	13.886	30.953	1.00	31.71	O
2054	O	HOH	530	-3.764	29.115	25.550	1.00	37.63	O
2055	O	HOH	531	31.831	9.097	66.550	1.00	36.20	O
2056	O	HOH	532	10.178	6.595	32.965	1.00	30.94	O

2057	O	HOH	533	-1.561	14.197	34.245	1.00	33.20	O
2058	O	HOH	534	0.476	12.154	62.160	1.00	39.93	O
2059	O	HOH	535	25.970	5.142	53.011	1.00	47.31	O
2060	O	HOH	536	8.695	5.045	44.801	1.00	38.39	O
2061	O	HOH	537	22.396	11.047	39.112	1.00	40.45	O
2062	O	HOH	538	13.975	29.983	22.553	1.00	36.21	O
2063	O	HOH	539	-6.673	18.195	37.122	1.00	36.41	O
2064	O	HOH	540	15.926	27.813	55.197	1.00	43.43	O
2065	O	HOH	541	21.922	29.786	26.625	1.00	39.42	O
2066	O	HOH	542	29.079	22.924	57.335	1.00	43.49	O
2067	O	HOH	543	-8.883	26.986	29.744	1.00	47.42	O
2068	O	HOH	544	-2.789	31.232	23.837	1.00	38.14	O
2069	O	HOH	545	15.578	33.329	45.128	1.00	39.44	O
2070	O	HOH	546	20.810	2.660	42.920	1.00	51.44	O
2071	O	HOH	547	27.448	25.982	58.310	1.00	43.04	O
2072	O	HOH	548	21.987	8.152	64.287	1.00	43.15	O
2073	O	HOH	549	14.435	13.091	64.840	1.00	35.87	O
2074	O	HOH	550	1.276	25.772	21.944	1.00	40.66	O
2075	O	HOH	551	14.102	6.513	31.763	1.00	43.70	O
2076	O	HOH	552	11.990	24.017	53.147	1.00	45.62	O
2077	O	HOH	553	3.481	24.236	20.666	1.00	35.69	O
2078	O	HOH	554	24.054	13.110	35.770	1.00	37.92	O
2079	O	HOH	556	6.857	37.182	44.351	1.00	49.60	O
2080	O	HOH	557	-8.644	30.901	30.925	1.00	41.21	O
2081	O	HOH	558	17.767	33.571	43.159	1.00	37.66	O
2082	O	HOH	559	16.954	26.537	23.238	1.00	51.77	O
2083	O	HOH	560	27.386	20.638	40.959	1.00	37.25	O
2084	O	HOH	561	31.418	10.182	50.496	1.00	47.27	O
2085	O	HOH	562	4.082	21.082	20.610	1.00	37.94	O
2086	O	HOH	563	14.064	10.706	58.224	1.00	42.75	O
2087	O	HOH	564	23.415	29.835	49.803	1.00	45.77	O
2088	O	HOH	565	14.533	11.393	24.395	1.00	36.60	O
2089	O	HOH	566	-0.868	36.798	40.025	1.00	52.17	O
2090	O	HOH	567	2.865	34.386	33.570	1.00	42.56	O
2091	O	HOH	568	-4.893	19.288	30.751	1.00	44.30	O
2092	O	HOH	569	30.643	14.674	61.949	1.00	43.28	O
2093	O	HOH	570	22.702	3.372	47.417	1.00	36.93	O
2094	O	HOH	571	13.379	35.172	44.109	1.00	47.38	O
2095	O	HOH	572	-1.138	20.698	22.966	1.00	53.61	O
2096	O	HOH	573	25.589	19.849	33.401	1.00	52.13	O
2097	O	HOH	574	23.893	13.360	32.579	1.00	45.26	O
2098	O	HOH	575	-7.367	18.485	31.944	1.00	48.23	O
2099	O	HOH	576	2.430	19.200	65.790	1.00	45.13	O
2100	O	HOH	577	20.048	32.028	44.907	1.00	46.82	O
2101	O	HOH	578	20.286	6.713	37.519	1.00	43.08	O
2102	O	HOH	579	25.879	5.448	50.403	1.00	48.82	O
2103	O	HOH	580	24.905	19.763	39.659	1.00	45.39	O
2104	O	HOH	581	2.341	14.233	26.082	1.00	50.76	O
2105	O	HOH	582	15.248	20.000	60.506	1.00	44.08	O
2106	O	HOH	583	22.695	7.038	37.715	1.00	46.55	O
2107	O	HOH	584	11.915	16.625	66.479	1.00	52.58	O
2108	O	HOH	585	20.145	35.730	35.936	1.00	46.90	O
2109	O	HOH	586	10.735	24.933	16.684	1.00	46.64	O
2110	O	HOH	587	1.182	9.495	61.830	1.00	55.88	O
2111	O	HOH	588	-3.993	16.527	51.745	1.00	43.33	O

2112	O	HOH	589	21.842	29.919	56.624	1.00	42.17	O
2113	O	HOH	590	3.602	25.520	44.494	1.00	50.24	O
2114	O	HOH	591	1.198	23.984	44.777	1.00	43.76	O
2115	O	HOH	592	13.208	27.713	54.123	1.00	59.17	O
2116	O	HOH	593	27.958	7.530	50.434	1.00	53.55	O
2117	O	HOH	594	22.594	3.510	64.140	1.00	45.66	O
2118	O	HOH	595	30.412	22.979	36.623	1.00	71.37	O
2119	O	HOH	596	10.560	15.906	20.574	1.00	50.32	O
2120	O	HOH	597	26.021	3.241	64.667	1.00	49.85	O
2121	O	HOH	598	19.853	9.062	62.967	1.00	56.45	O
2122	O	HOH	599	12.462	3.992	52.363	1.00	42.46	O
2123	O	HOH	600	6.152	35.657	28.721	1.00	46.87	O
2124	O	HOH	601	7.626	29.983	53.085	1.00	51.73	O
2125	O	HOH	602	11.547	23.591	57.064	1.00	51.07	O
2126	O	HOH	603	24.407	19.393	31.035	1.00	53.85	O
2127	O	HOH	604	12.538	23.006	18.706	1.00	50.11	O
2128	O	HOH	605	1.839	16.469	66.997	1.00	49.40	O
2129	O	HOH	606	1.378	19.964	21.070	1.00	48.97	O
2130	O	HOH	607	5.895	26.935	51.419	1.00	53.95	O
2131	O	HOH	608	13.122	33.698	19.464	1.00	52.90	O
2132	O	HOH	609	27.040	8.636	44.102	1.00	44.22	O
2133	O	HOH	610	18.833	30.775	55.879	1.00	54.75	O
2134	O	HOH	611	34.509	17.720	47.771	1.00	42.84	O
2135	O	HOH	612	18.356	32.644	25.579	1.00	42.52	O
2136	O	HOH	613	-2.259	16.235	28.804	1.00	56.71	O
2137	O	HOH	614	16.400	38.404	21.700	1.00	46.19	O
2138	O	HOH	615	9.340	39.540	19.060	1.00	51.44	O
2139	O	HOH	616	20.026	35.074	32.855	1.00	47.06	O
2140	O	HOH	617	31.604	8.486	59.428	1.00	47.99	O
2141	O	HOH	618	26.228	8.975	40.708	1.00	47.20	O
2142	O	HOH	619	0.460	15.378	28.064	1.00	50.21	O
2143	O	HOH	620	15.771	3.385	48.139	1.00	38.09	O
2144	O	HOH	621	25.135	17.914	42.644	1.00	60.05	O
2145	O	HOH	622	-2.286	29.197	21.618	1.00	53.99	O
2146	O	HOH	623	32.865	18.926	45.658	1.00	48.11	O
2147	O	HOH	624	17.116	13.333	25.240	1.00	52.60	O
2148	O	HOH	625	-2.809	17.978	56.255	1.00	53.36	O
2149	O	HOH	626	-3.647	7.885	56.347	1.00	63.91	O
2150	O	HOH	627	17.746	24.596	21.608	1.00	59.81	O
2151	O	HOH	628	28.368	5.841	47.861	1.00	66.08	O
2152	O	HOH	629	13.641	11.618	66.858	1.00	52.02	O
2153	O	HOH	630	8.052	20.893	16.742	1.00	53.91	O
2154	O	HOH	631	8.914	38.015	27.578	1.00	56.47	O
2155	O	HOH	632	9.081	13.482	19.627	1.00	57.14	O
2156	O	HOH	633	-4.343	24.969	37.694	1.00	51.08	O
2157	O	HOH	634	3.597	28.859	46.576	1.00	54.80	O
2158	O	HOH	635	27.905	21.432	28.373	1.00	59.49	O
2159	O	HOH	636	-4.252	18.337	25.491	1.00	47.50	O
2160	O	HOH	637	-2.808	23.046	51.839	1.00	49.04	O
2161	O	HOH	638	2.757	25.756	18.437	1.00	49.80	O
2162	O	HOH	639	15.470	7.390	63.803	1.00	52.42	O
2163	O	HOH	640	33.689	11.757	50.784	1.00	54.00	O
2164	O	HOH	641	6.223	13.352	20.927	1.00	49.77	O
2165	O	HOH	642	12.267	32.764	51.605	1.00	48.76	O
2166	O	HOH	644	25.211	3.585	48.391	1.00	49.75	O

2167	O	HOH	645	0.619	24.002	51.358	1.00	49.46	O
2168	O	HOH	646	12.270	22.627	60.617	1.00	63.88	O
2169	O	HOH	647	0.202	23.805	47.834	1.00	52.54	O
2170	O	HOH	648	15.471	8.169	23.816	1.00	54.49	O
2171	O	HOH	649	4.098	13.117	28.105	1.00	43.97	O
2172	O	HOH	650	16.032	4.857	59.064	1.00	55.67	O
2173	O	HOH	651	-5.591	11.911	55.960	1.00	63.35	O
2174	O	HOH	652	14.373	4.083	36.218	1.00	49.18	O
2175	O	HOH	653	11.138	5.501	59.825	1.00	51.19	O
2176	O	HOH	654	26.262	1.299	50.288	1.00	61.20	O
2177	O	HOH	655	4.067	20.751	67.111	1.00	51.75	O
2178	O	HOH	656	11.291	34.551	23.646	1.00	53.35	O
2179	O	HOH	657	2.505	33.743	45.342	1.00	58.29	O
2180	O	HOH	658	18.881	-0.886	43.452	1.00	60.82	O
2181	O	HOH	659	-1.930	13.191	62.255	1.00	65.05	O
2182	O	HOH	660	-3.587	12.153	34.625	1.00	51.24	O
2183	O	HOH	661	-2.064	26.008	58.110	1.00	58.94	O
2184	O	HOH	662	18.842	12.351	64.527	1.00	60.06	O
2185	O	HOH	663	30.991	26.420	51.105	1.00	54.69	O
2186	O	HOH	664	16.115	30.354	56.207	1.00	60.96	O
2187	O	HOH	665	36.596	19.242	55.988	1.00	55.83	O

It will be understood that various details of the invention can be changed without departing from the scope of the invention. Furthermore, the foregoing description is for the purpose of illustration only, and not for the purpose of limitation—the invention being defined by the claims.

CLAIMS

What is claimed is:

1. A substantially pure PXR ligand binding domain polypeptide in crystalline form.

2. The polypeptide of claim 1, wherein the crystalline form has lattice constants of a = 91.6 Å, b = 91.6 Å, c = 85.0 Å, α = 90°, β = 90°, γ = 90°.

3. The polypeptide of claim 1 or 2, wherein the crystalline form is a tetragonal crystalline form.

4. The polypeptide of claim 1 or 2, wherein the crystalline form has a space group of $P4_32_12$.

5. The polypeptide of claim 1 or 2, wherein the PXR ligand binding domain polypeptide has the amino acid sequence shown in SEQ ID NO: 4.

6. The polypeptide of claim 1 or 2, wherein the PXR ligand binding domain polypeptide is in complex with a ligand.

7. The polypeptide of claim 6, wherein the ligand is a drug.

8. The polypeptide of claim 7, wherein the drug is SR12813.

9. The polypeptide of claim 1 or 2, wherein the PXR ligand binding domain has a crystalline structure further characterized by the coordinates corresponding to Table 4.

10. The polypeptide of claim 1 or 2, wherein the crystalline form contains one PXR ligand binding domain polypeptide in the asymmetric unit.

11. The polypeptide of claim 1 or 2, wherein the crystalline form is such that the three-dimensional structure of the crystallized hPXR ligand binding domain polypeptide can be determined to a resolution of about 2.75 Å or better.

12. The polypeptide of claim 1 or 2, wherein the crystalline form contains one or more atoms having an atomic weight of 40 grams/mol or greater.

13. A method for determining the three-dimensional structure of a crystallized hPXR ligand binding domain polypeptide to a resolution of about 3.0 Å or better, the method comprising:

(a) crystallizing a hPXR ligand binding domain polypeptide; and

(b) analyzing the hPXR ligand binding domain polypeptide to determine the three-dimensional structure of the crystallized hPXR ligand binding domain polypeptide, whereby the three-dimensional structure of a crystallized hPXR ligand binding domain polypeptide is determined to a resolution of about 2.75 Å or better.

14. The method of claim 13, wherein the analyzing is by X-ray diffraction.

15. The method of claim 13, wherein the crystallization is accomplished by the hanging drop vapor diffusion method, and wherein the PXR ligand binding domain is mixed with an equal volume of reservoir.

16. The method of claim 15, wherein the reservoir comprises 250 mM NaCl, 20 mM Tris-HCl pH 7.8, 5% glycerol (v/v), 5 mM DTT, 2.5 mM EDTA.

17. The method of claim 15, wherein the reservoir 50 mM imidazole, pH 7.4, 10% 2-propanol (v/v).

18. A method of generating a crystallized PXR ligand binding domain polypeptide, the method comprising:

(a) incubating a solution comprising a PXR ligand binding domain with an equal volume of reservoir; and

(b) crystallizing the PXR ligand binding domain polypeptide using the hanging drop method, whereby a crystallized PXR ligand binding domain polypeptide is generated.

19. A crystallized PXR ligand binding domain polypeptide produced by the method of claim 18.

20. A method of designing a modulator of a PXR polypeptide, the method comprising:

(a) designing a potential modulator of a PXR polypeptide that will form bonds with amino acids in a ligand binding site based upon a crystalline structure of a PXR ligand binding domain polypeptide;

(b) synthesizing the modulator; and

(c) determining whether the potential modulator modulates the activity of the PXR polypeptide, whereby a modulator of a PXR polypeptide is designed.

21. A method of designing a modulator that selectively modulates the activity of a human PXR polypeptide, the method comprising:

(a) obtaining a crystalline form of a human PXR ligand binding domain polypeptide;

(b) evaluating the three-dimensional structure of the crystallized human PXR ligand binding domain polypeptide; and

(c) synthesizing a potential modulator based on the three-dimensional crystal structure of the crystallized human PXR ligand binding domain polypeptide, whereby a modulator that selectively modulates the activity of a human PXR polypeptide is designed.

22. The method of claim 21, wherein the method further comprises contacting a human PXR ligand binding domain polypeptide with the potential modulator; and assaying the human PXR ligand binding domain polypeptide for binding of the potential modulator, for a change in activity of the human PXR ligand binding domain polypeptide, or both.

23. The method of claim 21, wherein the crystalline form is in tetragonal form.

24. The method of claim 23, wherein the crystalline form is such that the three-dimensional structure of the crystallized human PXR ligand binding domain polypeptide can be determined to a resolution of about 2.75 Å or better.

25. A method for identifying a PXR modulator, the method comprising:

(a) providing atomic coordinates of a PXR ligand binding domain to a computerized modeling system; and

(b) modeling a ligand that fits spatially into a binding cavity or on the surface of the PXR ligand binding domain, whereby a PXR modulator is identified.

26. The method of claim 25, wherein the method further comprises identifying in an assay for PXR-mediated activity a modeled ligand that increases or decreases the activity of the PXR.

27. A method of identifying a PXR modulator that selectively modulates the activity of a PXR polypeptide compared to other polypeptides, the method comprising:

(a) providing atomic coordinates of a PXR ligand binding domain to a computerized modeling system; and

(b) modeling a ligand that fits spatially into a binding cavity or on the surface of a PXR ligand binding domain and that interacts with conformationally constrained residues of a PXR that are conserved among PXR orthologs and isoforms, whereby a PXR modulator is identified.

28. The method of claim 27, wherein the method further comprises identifying in a biological assay for PXR–mediated activity a modeled ligand that selectively binds to the PXR ligand binding domain and increases or decreases the activity of the PXR.

29. A method of designing a modulator of a PXR polypeptide, the method comprising:

(a) selecting a candidate PXR ligand;

(b) determining which amino acid or amino acids of a PXR polypeptide interact with the ligand using a three-dimensional model of a crystallized protein comprising a PXR LBD;

(c) identifying in a biological assay for PXR activity a degree to which the ligand modulates the activity of the PXR polypeptide;

(d) selecting a chemical modification of the ligand wherein the interaction between the amino acids of the PXR polypeptide and the ligand is predicted to be modulated by the chemical modification;

(e) performing the chemical modification on the ligand to form a modified ligand;

(f) contacting the modified ligand with the PXR polypeptide;

(g) identifying in a biological assay for PXR activity a degree to

which the modified ligand modulates the biological activity of the PXR polypeptide; and

(h) comparing the biological activity of the PXR polypeptide in the presence of modified ligand with the biological activity of the PXR polypeptide in the presence of the unmodified ligand, whereby a modulator of a PXR polypeptide is designed.

30. The method of claim 29, wherein the PXR polypeptide is a human PXR polypeptide.

31. The method of claim 29, wherein the three-dimensional model of a crystallized protein is a human PXR LBD polypeptide with a bound ligand.

32. The method of claim 31, wherein the ligand is a drug.

33. The method of claim 32, wherein the drug is SR12813.

34. The method of claim 29, wherein the method further comprises repeating steps (a) through (f), if the biological activity of the PXR polypeptide in the presence of the modified ligand varies from the biological activity of the PXR polypeptide in the presence of the unmodified ligand.

35. An assay method for identifying a compound that inhibits binding of a ligand to a PXR polypeptide, the assay method comprising:

(a) designing a test inhibitor compound capable of modulating PXR activity, based on the atomic coordinates of a PXR ligand binding domain;

(b) synthesizing the test inhibitor compound;

(c) incubating a PXR polypeptide with a ligand in the presence of a test inhibitor compound;

(d) determining an amount of ligand that is bound to the PXR polypeptide, wherein decreased binding of ligand to the PXR

protein in the presence of the test inhibitor compound relative to binding of ligand in the absence of the test inhibitor compound is indicative of inhibition; and

(e) identifying the test compound as an inhibitor of ligand binding if decreased ligand binding is observed, whereby a compound that inhibits binding of a ligand to a PXR polypeptide is identified.

36. The method of claim 35, wherein the ligand is a drug.

37. The method of claim 36, wherein the drug is SR12813.

38. A method of evaluating a candidate therapeutic agent in humans using a mouse model system, the method comprising:

(a) providing atomic coordinates of a human PXR ligand binding domain to a computerized modeling system;

(b) modeling a candidate therapeutic agent that fits spatially into a binding cavity or on the surface of a human PXR ligand binding domain;

(c) providing a mouse PXR polypeptide;

(d) selecting one or more mutations to be introduced into an amino acid sequence of the mouse PXR polypeptide, the mutations being selected so as to alter the mouse PXR polypeptide to be similar to a human PXR polypeptide;

(e) providing a mutant mouse PXR polypeptide comprising the one or more mutations selected in step (d);

(f) contacting the candidate therapeutic agent modeled in step (b) with the mutant mouse PXR polypeptide;

(f) determining an effect of the candidate therapeutic agent on the mutant mouse PXR polypeptide; and

(g) evaluating the potential of the candidate therapeutic agent for use in humans based on the effect of the candidate therapeutic agent on the mutant mouse PXR polypeptide, whereby the

candidate therapeutic agent in humans using a mouse model system is evaluated.

39. The method of claim 38, further comprising repeating steps (a) through (h) a desired number of times.

40. The method of claim 38, wherein the one or more mutations is selected from the group consisting of an arginine to leucine substitution at residue 203 of the mouse PXR polypeptide, a protein to serine substitution at residue 205 of the mouse PXR polypeptide, a glutamine to histidine substitution at residue 404 of the mouse PXR polypeptide and a glutamine to arginine substitution at residue 407 of the mouse PXR polypeptide.

41. The method of claim 38, wherein the mouse PXR polypeptide comprises a mouse PXR ligand binding domain (LBD).

42. The method of claim 38, wherein the mutant mouse PXR polypeptide is expressed by:

(a) introducing the selected mutation into the mouse PXR DNA sequence, thereby forming a mutant mouse PXR DNA sequence;

(b) transforming cells with the mutant mouse PXR DNA sequence; and

(c) expressing the mutant mouse PXR DNA sequence, thereby forming a mutant mouse PXR polypeptide.

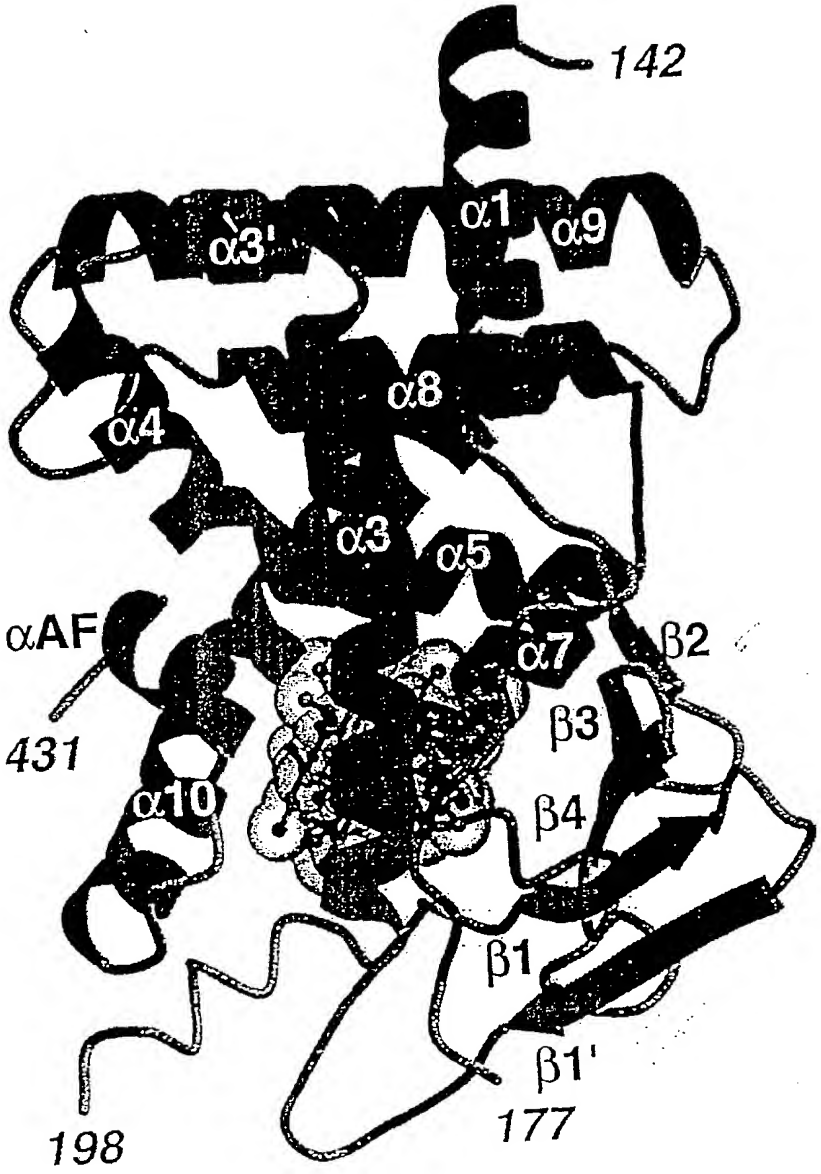
43. The method of claim 42, wherein the mutant mouse PXR polypeptide is expressed in cells.

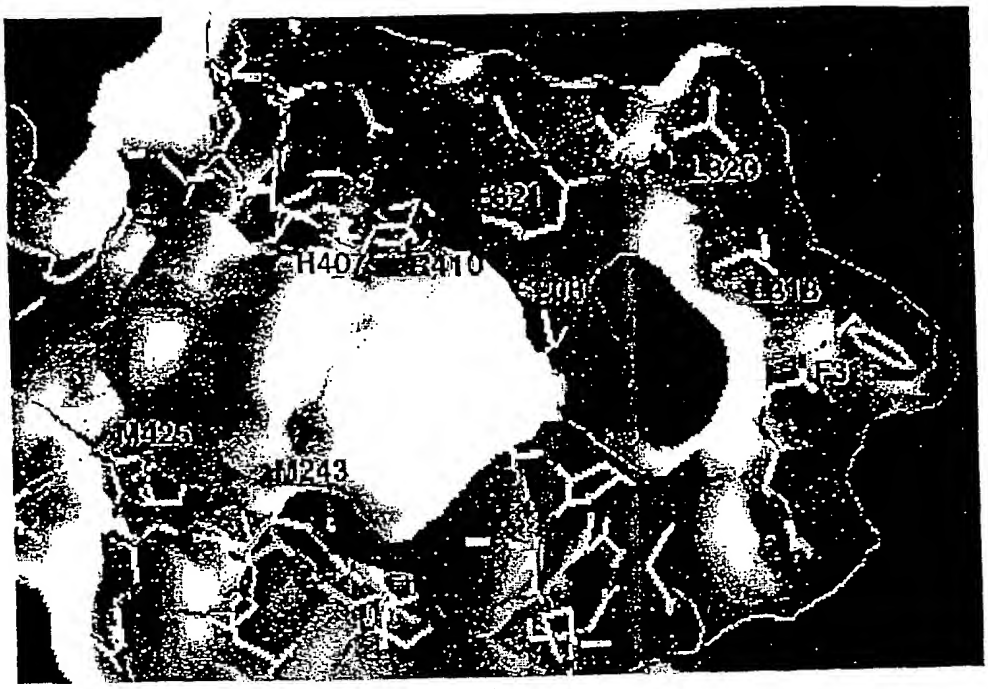
44. The method of claim 43, wherein the cells are CV-1 cells.

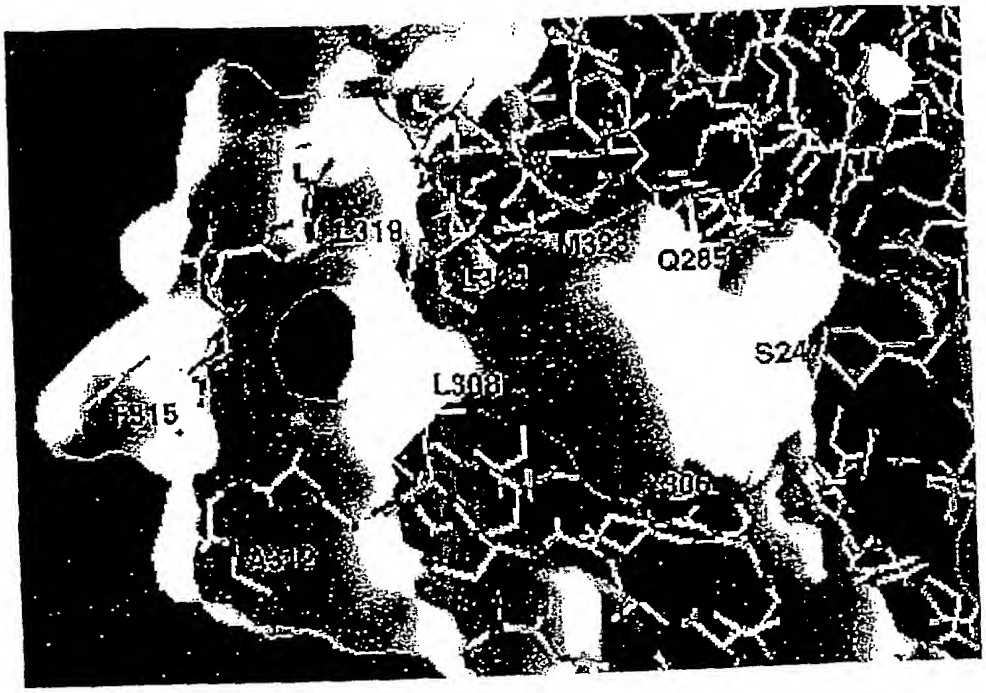
45. The method of claim 42, wherein the mutant mouse PXR DNA sequence is cotransformed with a reporter gene.

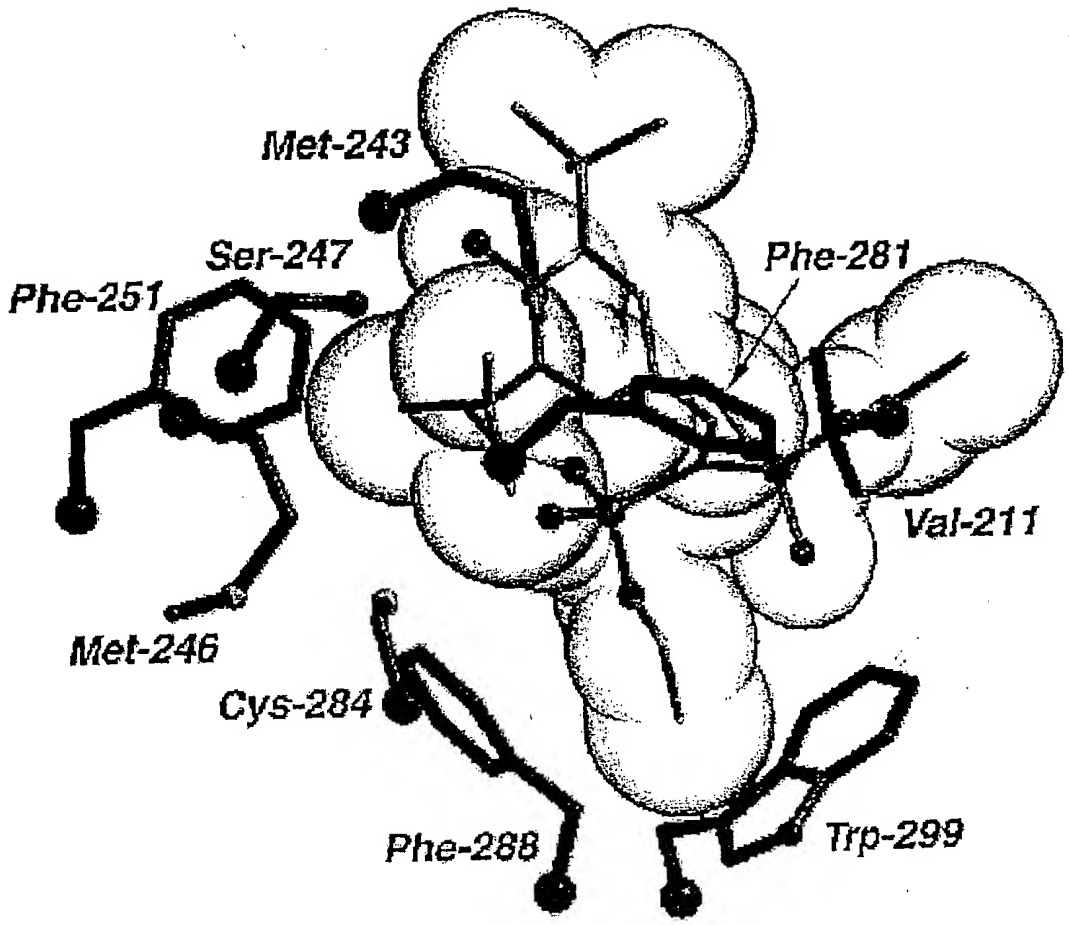
46. The method of claim 45, wherein the reporter gene is an XREM-CYP3A4-LUC reporter.

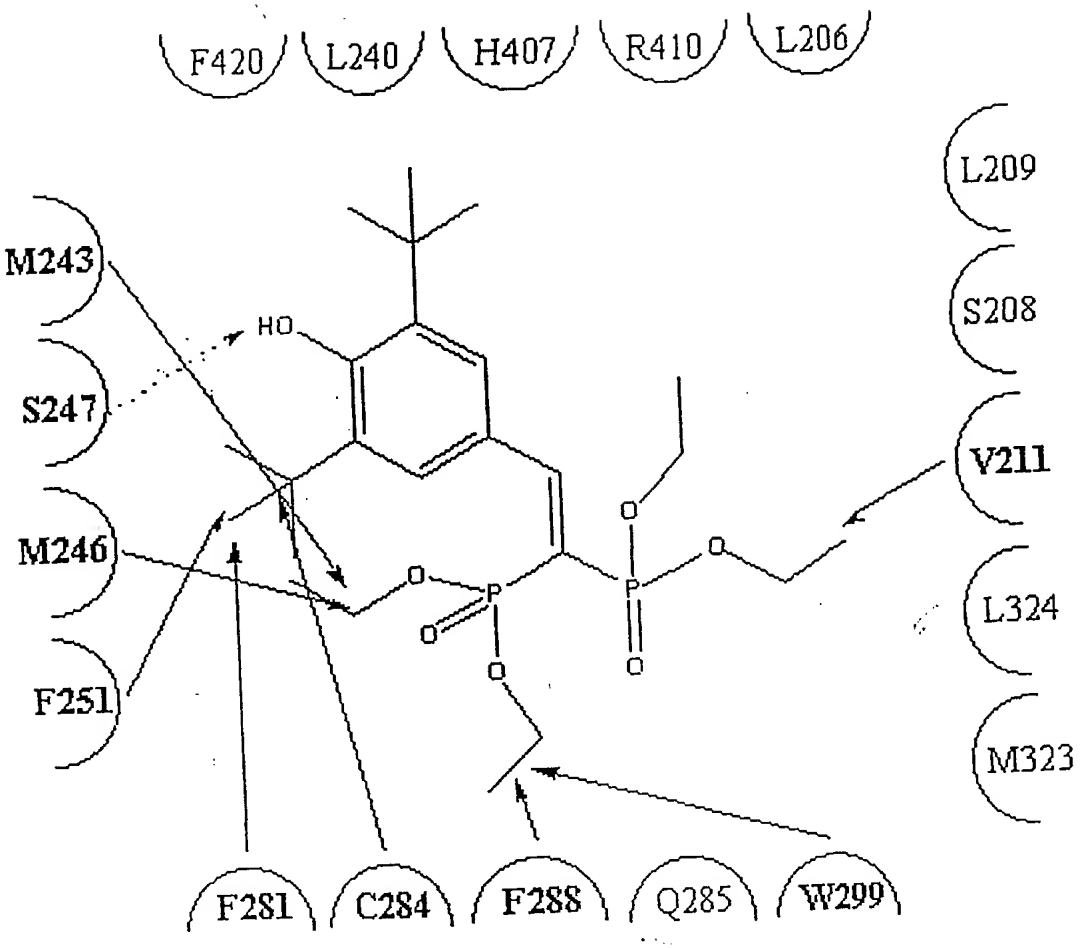
47. The method of claim 45, wherein the mouse PXR polypeptide comprises a mouse PXR ligand binding domain (LBD).

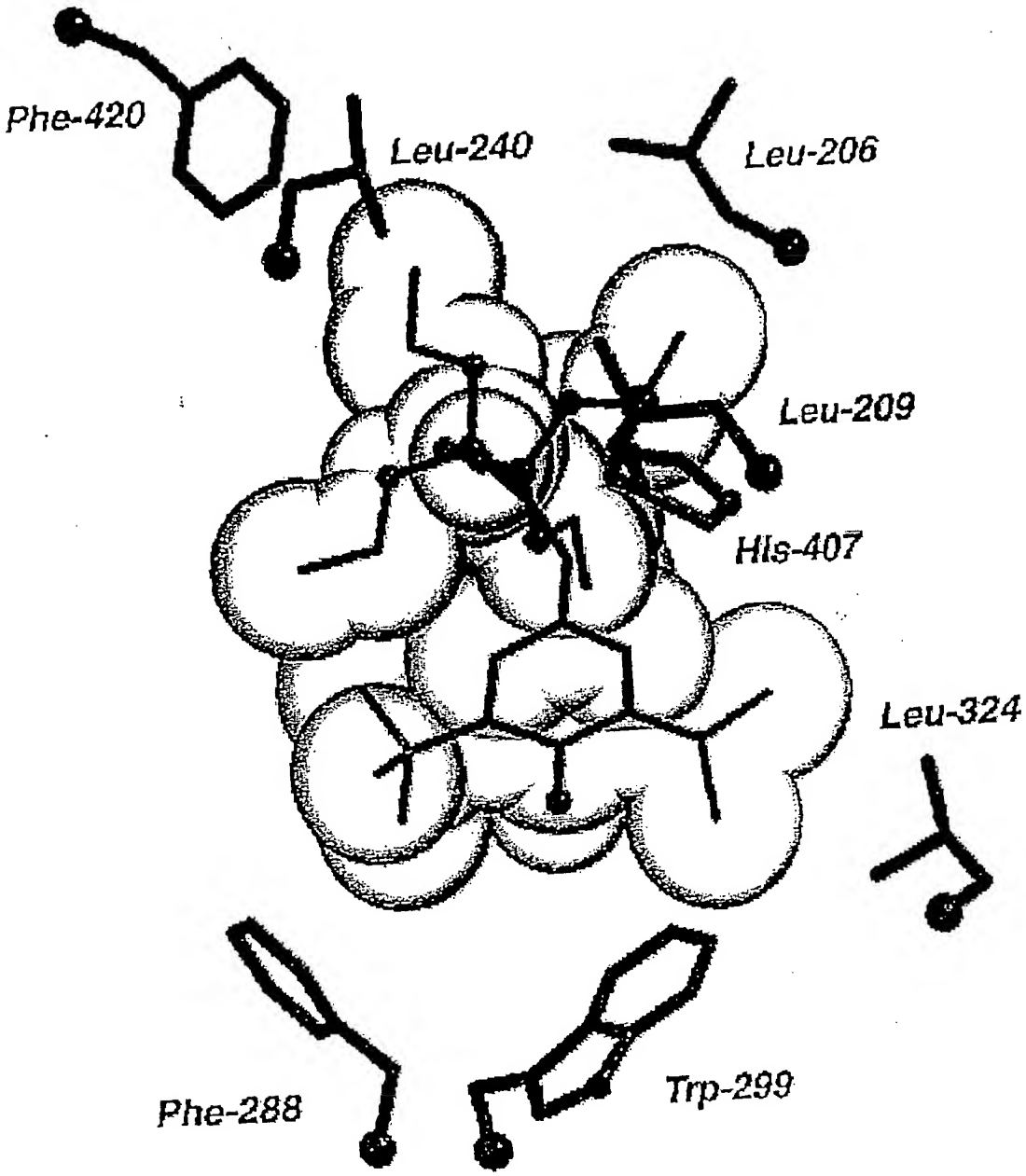

142
α1
α9
α3'
α8
α4
α3
α5
αAF
α7
β2
431
β3
α10
β4
β1
β1'
177
198

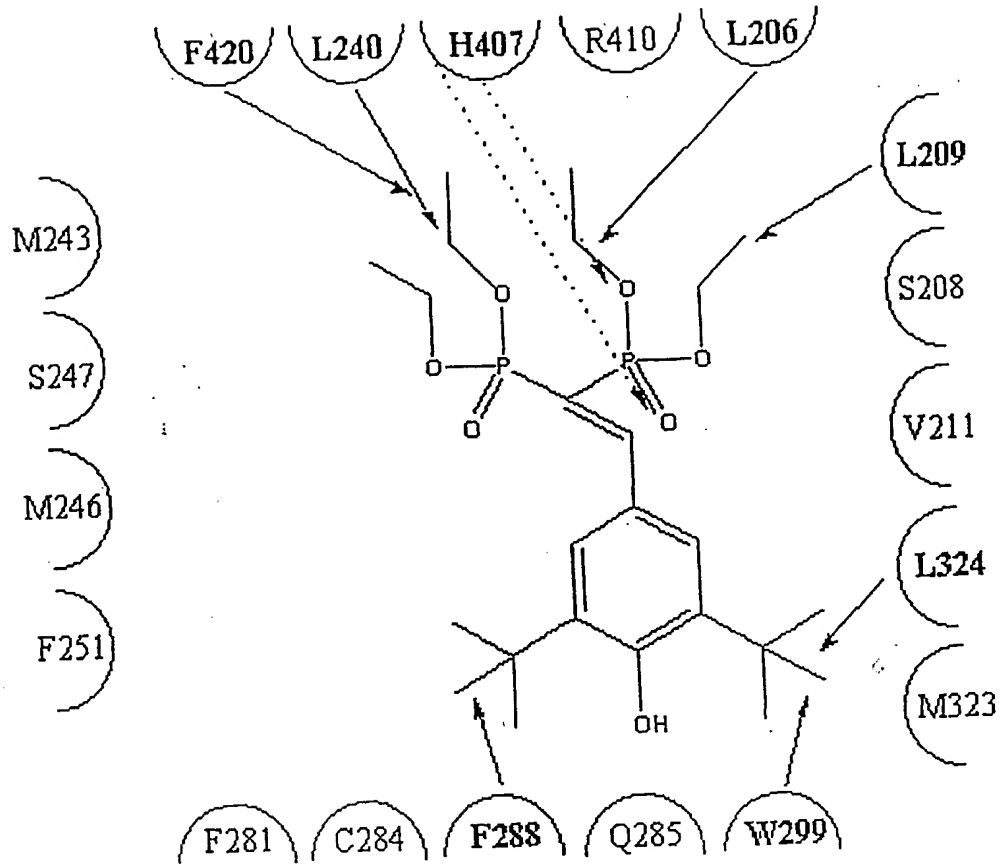

L320
H407
H410
L318
M425
M243

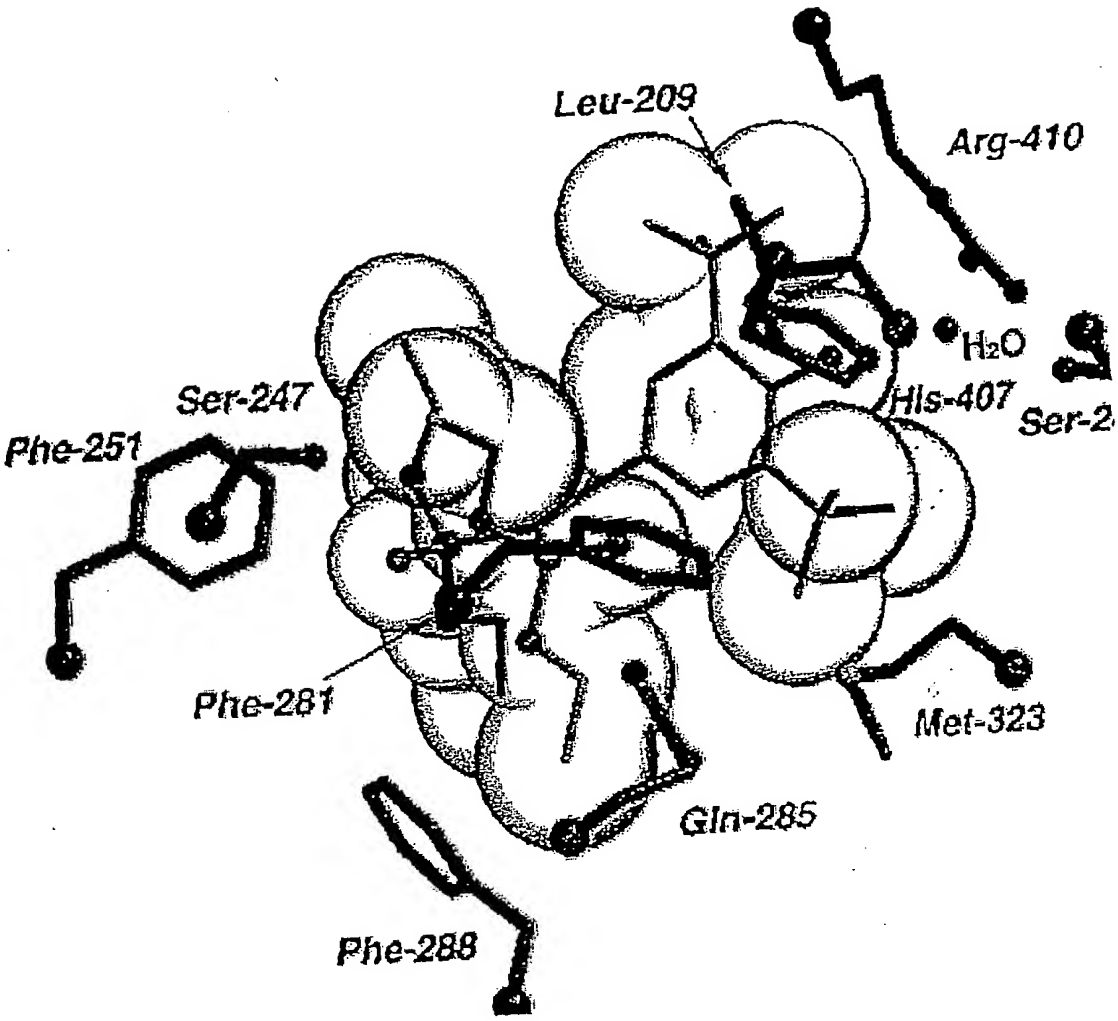

Q285
L308

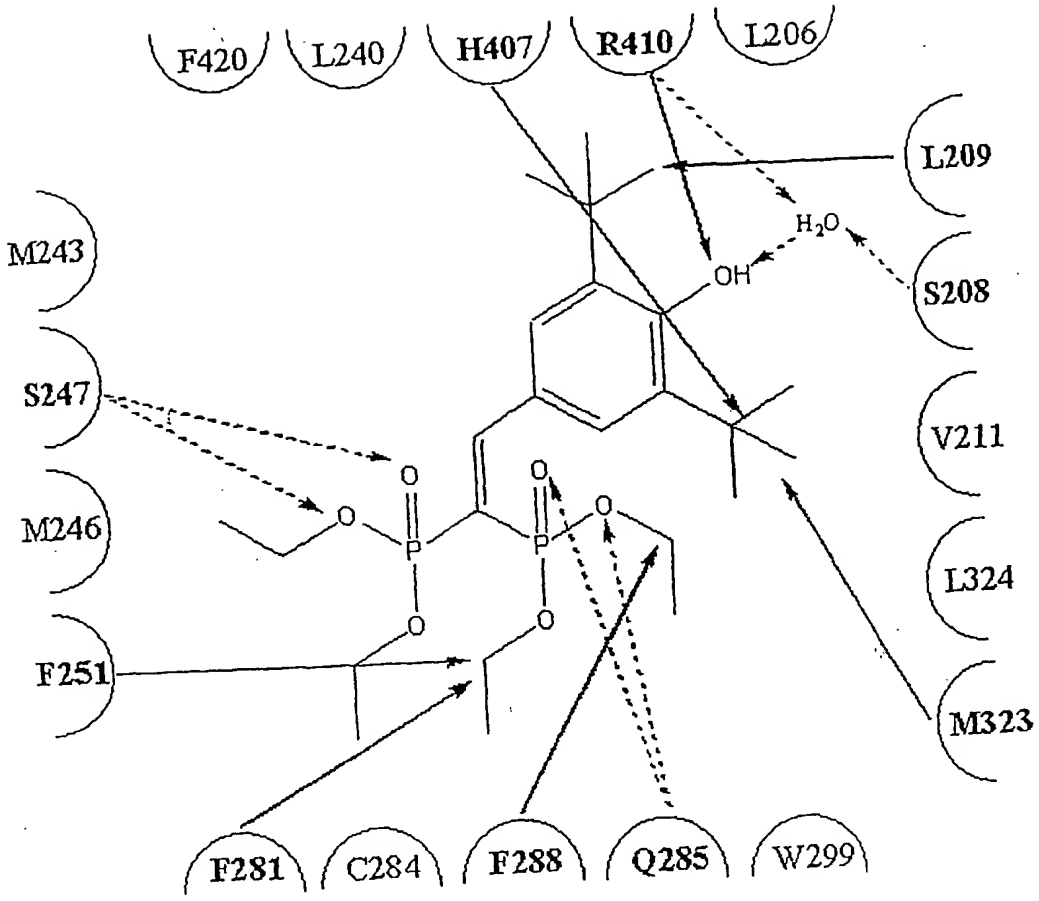

Met-243
Ser-247
Phe-251
Phe-281
Val-211
Met-246
Cys-284
Phe-288
Trp-299

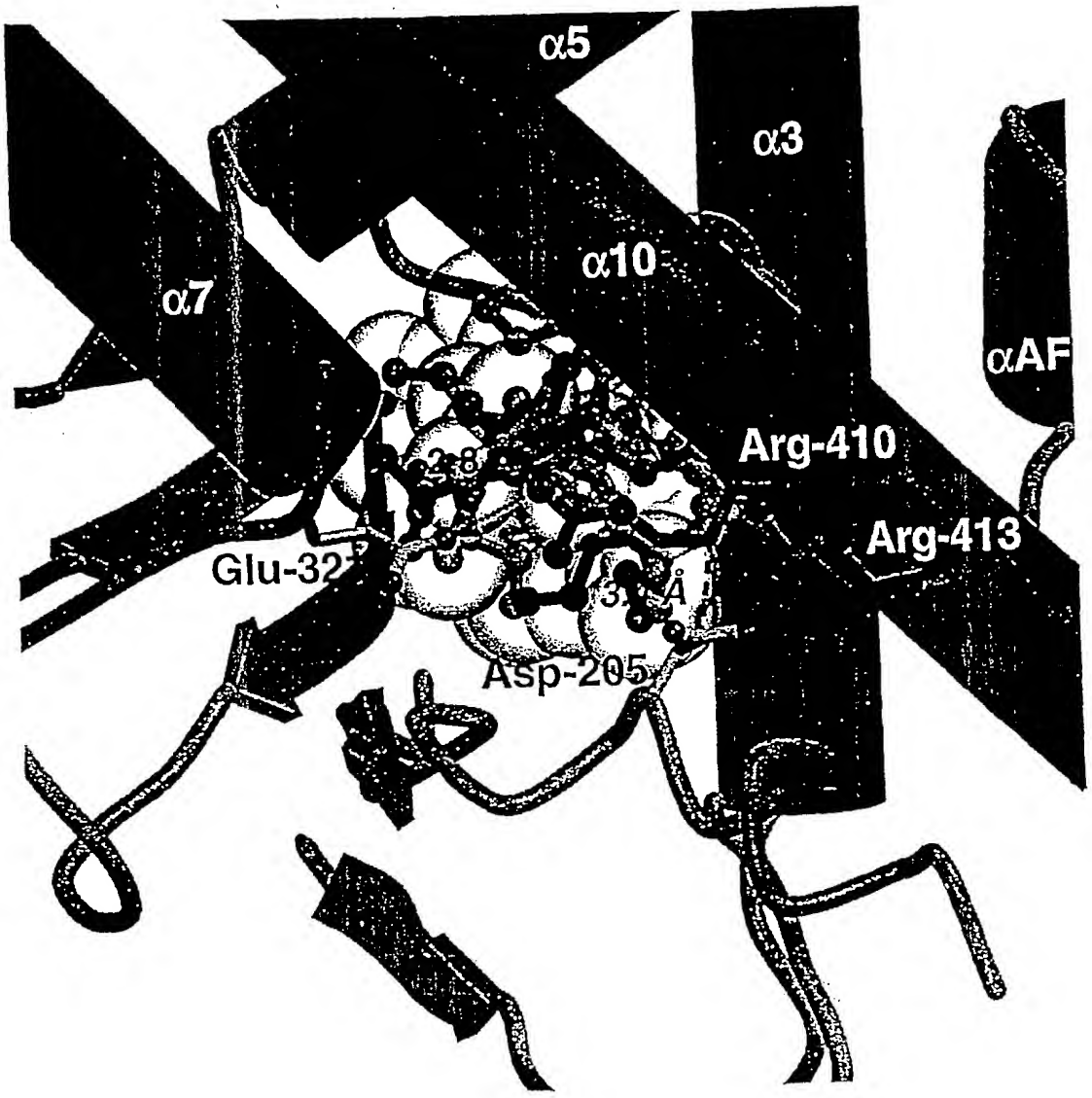

F420
L240
H407
R410
L206
L209
S208
V211
L324
M323
M243
S247
M246
F251
HO
O
P
O
O
O
O
P
O
O
F281
C284
F288
Q285
W299

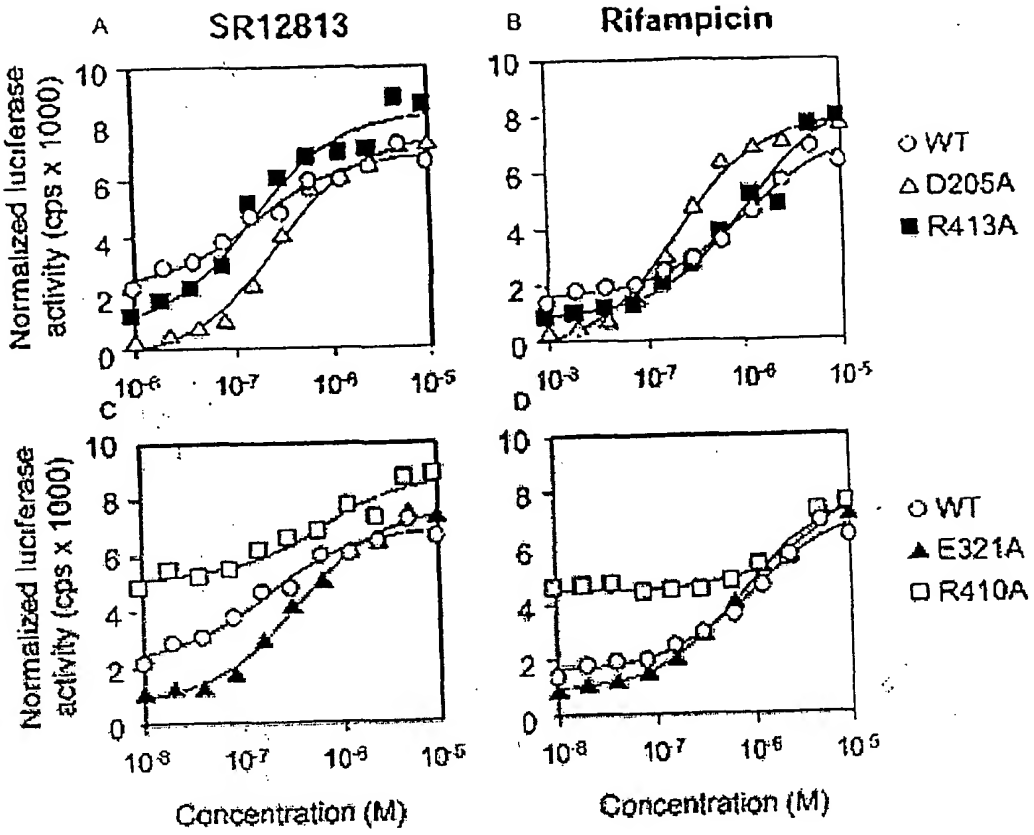

Phe-420
Leu-240
Leu-206
Leu-209
His-407
Leu-324
Phe-288
Trp-299

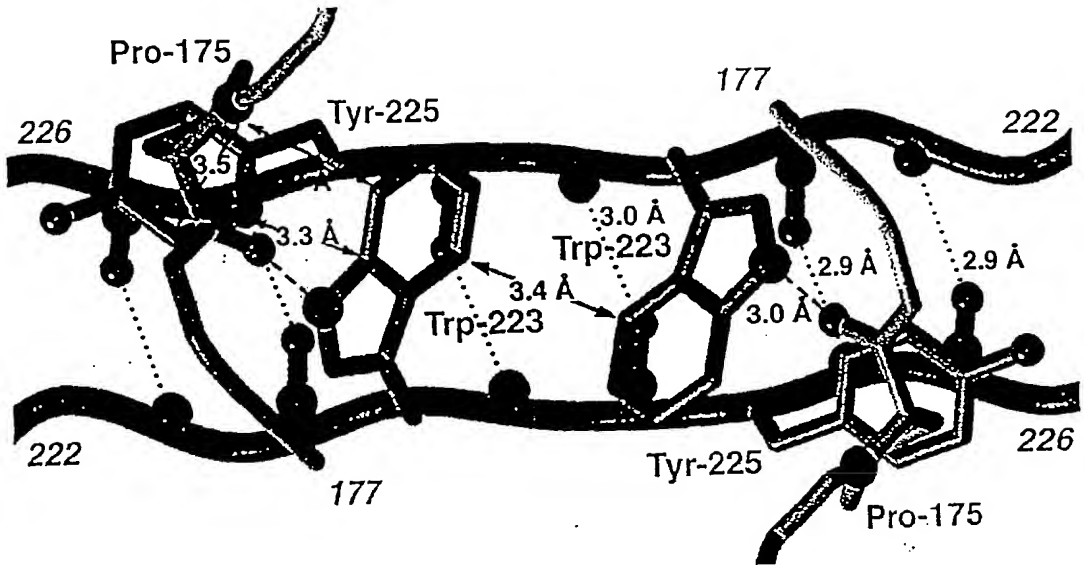

F420
L240
H407
R410
L206
L209
M243
S208
S247
V211
M246
L324
F251
M323
OH
F281
C284
F288
Q285
W299

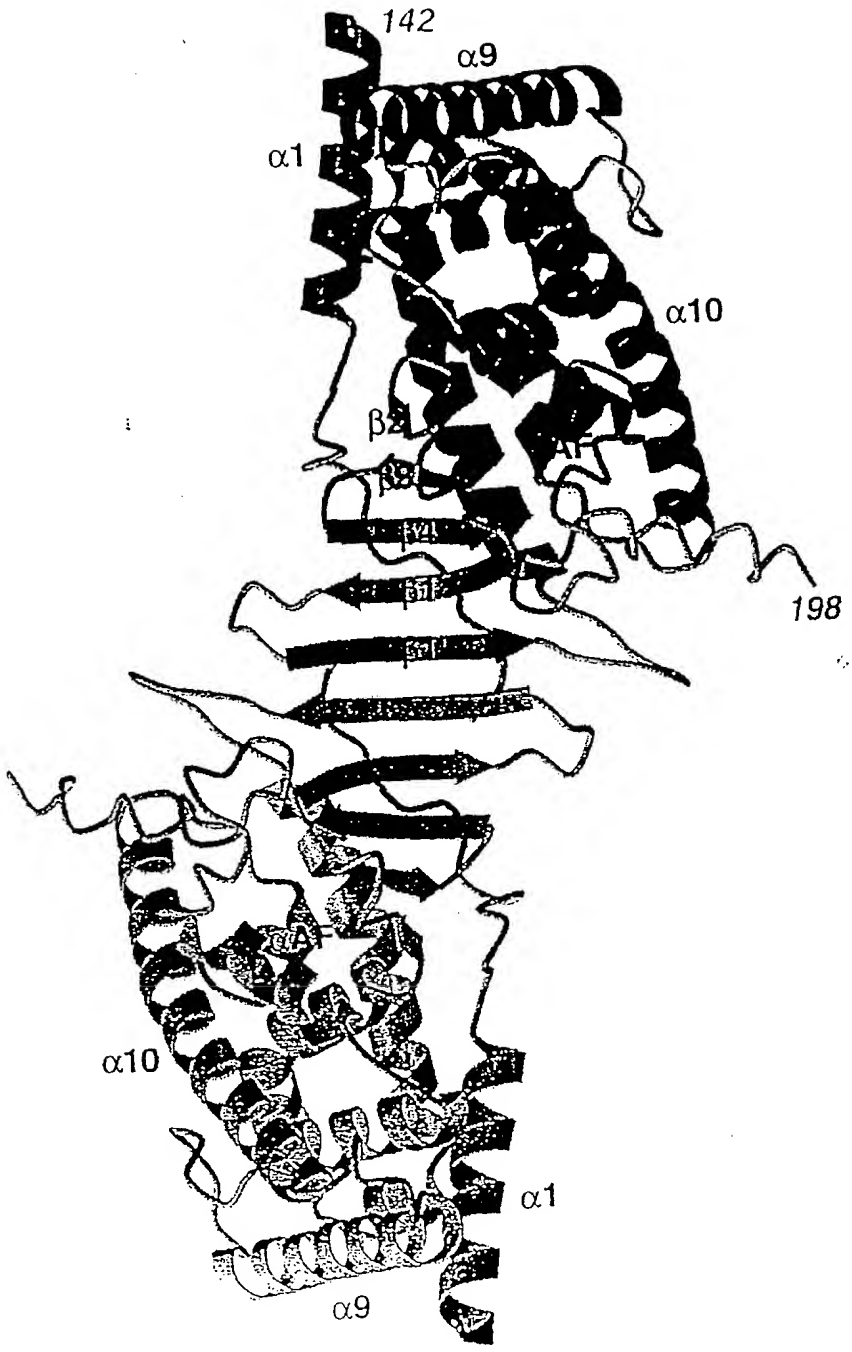

Leu-209
Arg-410
H2O
Ser-247
His-407
Ser-2
Phe-251
Phe-281
Met-323
Gln-285
Phe-288

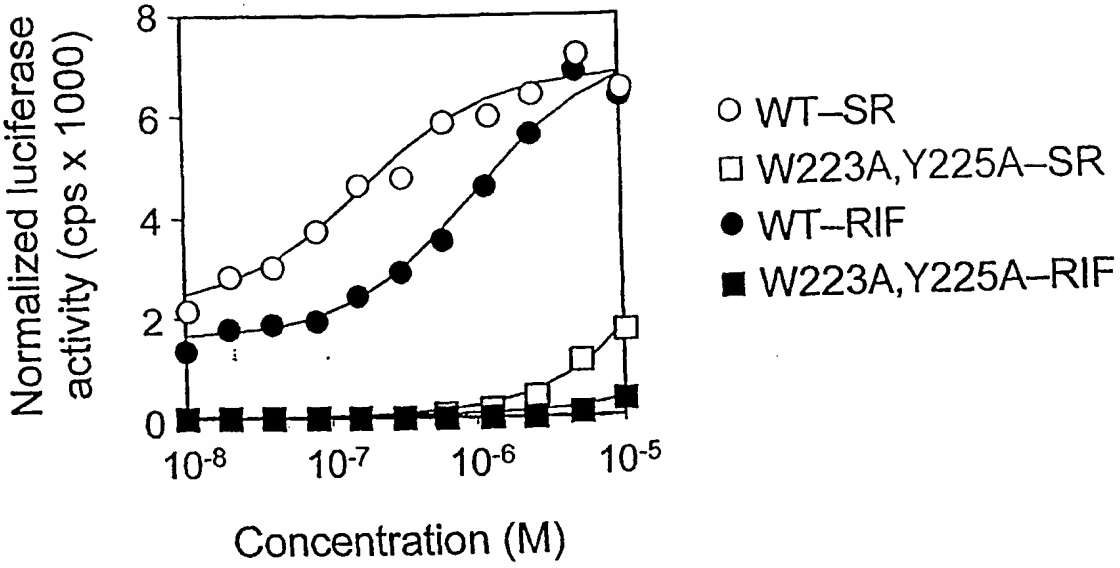

F420
L240
H407
R410
L206
L209
M243
H_2O
OH
S208
S247
V211
M246
L324
F251
M323
F281
C284
F288
Q285
W299

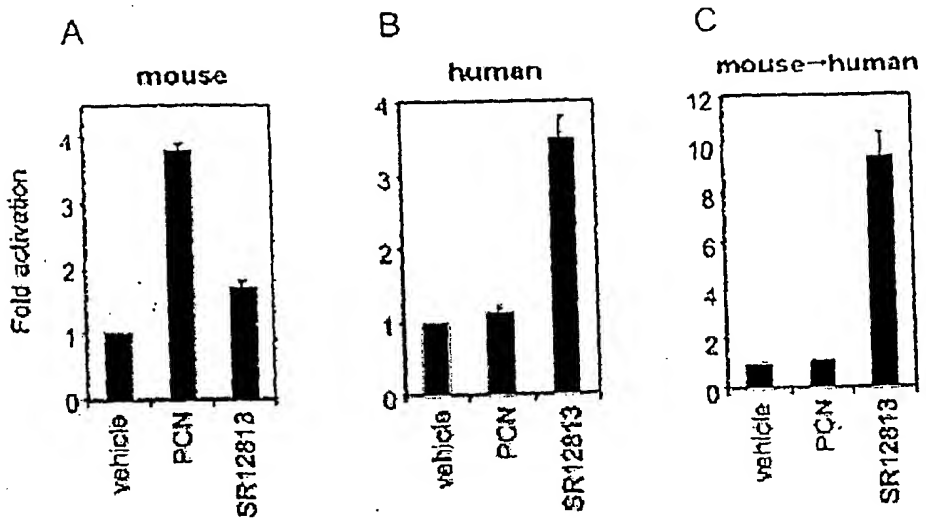

α5
α3
α10
α7
αAF
Arg-410
Arg-413
Glu-32
Asp-205


A
SR12813
B
Rifampicin
Normalized luciferase activity (cps x 1000)
WT
D205A
R413A
C
D
WT
E321A
R410A
Concentration (M)
Concentration (M)


Pro-175
Tyr-225
226
3.5
3.3 Å
177
222
3.0 Å
Trp-223
2.9 Å
2.9 Å
3.4 Å
Trp-223
3.0 Å
222
177
Tyr-225
226
Pro-175


142
α9
α1
α10
β2
β3
198
α10
α1
α9


Normalized luciferase activity (cps x 1000)
8
6
4
2
0
10-8
10-7
10-6
10-5
Concentration (M)
WT–SR
W223A,Y225A–SR
WT–RIF
W223A,Y225A–RIF


A
mouse
Fold activation
0
1
2
3
4
vehicle
PCN
SR12813
B
human
0
1
2
3
4
vehicle
PCN
SR12813
C
mouse→human
0
2
4
6
8
10
12
vehicle
PCN
SR12813

SEQUENCE LISTING

<110> SmithKline Beecham Corporation

Redinbo, Matthew

Watkins, Ryan

Wisely, George

Williams, Shawn

Kliewer, Steven

<120> Crystallized Human Xenobiotic Nuclear Receptor PXR/SXR Ligand Binding Domain Polypeptide and Screening Methods Employing Same

<130> PU4460

<160> 13

<170> PatentIn version 3.0

<210> 1

<211> 2146

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (304)..(1608)

<300>

<301> Lehmann et al.

<302> The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug

interactions

<303> Journal of Clinical Investigation

<304> 102

<305> 5

<306> 1016-23

<307> 1998

<308> Swiss-Prot/075469; GenBank/AF061056

<309> 1999-01-20

<313> (1)..(2146)

<400> 1

```
tgaaatatag gtgagagaca agattgtctc atatccgggg aaatcataac ctatgactag     60

gacgggaaga ggaagcactg cctttacttc agtgggaatc tcggcctcag cctgcaagcc    120

aagtgttcac agtgagaaaa gcaagagaat aagctaatac tcctgtcctg aacaaggcag    180

cggctccttg gtaaagctac tccttgatcg atcctttgca ccggattgtt caaagtggac    240

cccaggggag aagtcggagc aaagaactta ccaccaagca gtccaagagg cccagaagca    300

aac ctg gag gtg aga ccc aaa gaa agc tgg aac cat gct gac ttt gta      348
    Leu Glu Val Arg Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val
    1               5                  10                  15

cac tgt gag gac aca gag tct gtt cct gga aag ccc agt gtc aac gca      396
His Cys Glu Asp Thr Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala
                20                  25                  30

gat gag gaa gtc gga ggt ccc caa atc tgc cgt gta tgt ggg gac aag      444
Asp Glu Glu Val Gly Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys
            35                  40                  45

gcc act ggc tat cac ttc aat gtc atg aca tgt gaa gga tgc aag ggc      492
Ala Thr Gly Tyr His Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly
        50                  55                  60

ttt ttc agg agg gcc atg aaa cgc aac gcc cgg ctg agg tgc ccc ttc      540
Phe Phe Arg Arg Ala Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe
    65                  70                  75

cgg aag ggc gcc tgc gag atc acc cgg aag acc cgg cga cag tgc cag      588
Arg Lys Gly Ala Cys Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln
80                  85                  90                  95

gcc tgc cgc ctg cgc aag tgc ctg gag agc ggc atg aag aag gag atg      636
Ala Cys Arg Leu Arg Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met
                100                 105                 110

atc atg tcc gac gag gcc gtg gag gag agg cgg gcc ttg atc aag cgg      684
Ile Met Ser Asp Glu Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg
            115                 120                 125

aag aaa agt gaa cgg aca ggg act cag cca ctg gga gtg cag ggg ctg      732
Lys Lys Ser Glu Arg Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu
        130                 135                 140

aca gag gag cag cgg atg atg atc agg gag ctg atg gac gct cag atg      780
Thr Glu Glu Gln Arg Met Met Ile Arg Glu Leu Met Asp Ala Gln Met
    145                 150                 155
```

```
aaa acc ttt gac act acc ttc tcc cat ttc aag aat ttc cgg ctg cca      828
Lys Thr Phe Asp Thr Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro
160                 165                 170                 175

ggg gtg ctt agc agt ggc tgc gag ttg cca gag tct ctg cag gcc cca      876
Gly Val Leu Ser Ser Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro
                180                 185                 190

tcg agg gaa gaa gct gcc aag tgg agc cag gtc cgg aaa gat ctg tgc      924
Ser Arg Glu Glu Ala Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys
            195                 200                 205

tct ttg aag gtc tct ctg cag ctg cgg ggg gag gat ggc agt gtc tgg      972
Ser Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp
        210                 215                 220

aac tac aaa ccc cca gcc gac agt ggc ggg aaa gag atc ttc tcc ctg     1020
Asn Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu
    225                 230                 235

ctg ccc cac atg gct gac atg tca acc tac atg ttc aaa ggc atc atc     1068
Leu Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile
240                 245                 250                 255

agc ttt gcc aaa gtc atc tcc tac ttc agg gac ttg ccc atc gag gac     1116
Ser Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp
                260                 265                 270

cag atc tcc ctg ctg aag ggg gcc gct ttc gag ctg tgt caa ctg aga     1164
Gln Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg
            275                 280                 285

ttc aac aca gtg ttc aac gcg gag act gga acc tgg gag tgt ggc cgg     1212
Phe Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg
        290                 295                 300

ctg tcc tac tgc ttg gaa gac act gca ggt ggc ttc cag caa ctt cta     1260
Leu Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu
    305                 310                 315

ctg gag ccc atg ctg aaa ttc cac tac atg ctg aag aag ctg cag ctg     1308
Leu Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu
320                 325                 330                 335

cat gag gag gag tat gtg ctg atg cag gcc atc tcc ctc ttc tcc cca     1356
His Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro
                340                 345                 350

gac cgc cca ggt gtg ctg cag cac cgc gtg gtg gac cag ctg cag gag     1404
Asp Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln Leu Gln Glu
            355                 360                 365

caa ttc gcc att act ctg aag tcc tac att gaa tgc aat cgg ccc cag     1452
Gln Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln
        370                 375                 380

cct gct cat agg ttc ttg ttc ctg aag atc atg gct atg ctc acc gag     1500
Pro Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu
    385                 390                 395

ctc cgc agc atc aat gct cag cac acc cag cgg ctg ctg cgc atc cag     1548
Leu Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln
400                 405                 410                 415

gac ata cac ccc ttt gct acg ccc ctc atg cag gag ttg ttc ggc atc     1596
Asp Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile
                420                 425                 430

aca ggt agc tga gcggctgccc ttgggtgaca cctccgagag gcagccagac        1648
Thr Gly Ser

ccagagccct ctgagccgcc actcccgggc caagacagat ggacactgcc aagagccgac  1708
```

aatgccctgc tggcctgtct ccctagggaa ttcctgctat gacagctggc tagcattcct 1768

caggaaggac atgggtgccc cccacccccа gttcagtctg tagggagtga agccacagac 1828

tcttacgtgg agagtgcact gacctgtagg tcaggaccat cagagaggca aggttgccct 1888

ttccttttaa aaggccctgt ggtctgggga gaaatccctc agatcccact aaagtgtcaa 1948

ggtgtggaag ggaccaagcg accaaggata ggccatctgg ggtctatgcc cacataccca 2008

cgtttgttcg cttcctgagt cttttcattg ctacctctaa tagtcctgtc tcccacttcc 2068

cactcgttcc cctcctcttc cgagctgctt tgtgggctcc aggcctgtac tcatcggcag 2128

gtgcatgagt atctgtgg 2146

<210> 2

<211> 434

<212> PRT

<213> Homo sapiens

<400> 2

Leu Glu Val Arg Pro Lys Glu Ser Trp Asn His Ala Asp Phe Val His
1 5 10 15

Cys Glu Asp Thr Glu Ser Val Pro Gly Lys Pro Ser Val Asn Ala Asp
20 25 30

Glu Glu Val Gly Gly Pro Gln Ile Cys Arg Val Cys Gly Asp Lys Ala
35 40 45

Thr Gly Tyr His Phe Asn Val Met Thr Cys Glu Gly Cys Lys Gly Phe
50 55 60

Phe Arg Arg Ala Met Lys Arg Asn Ala Arg Leu Arg Cys Pro Phe Arg
65 70 75 80

Lys Gly Ala Cys Glu Ile Thr Arg Lys Thr Arg Arg Gln Cys Gln Ala
85 90 95

Cys Arg Leu Arg Lys Cys Leu Glu Ser Gly Met Lys Lys Glu Met Ile
100 105 110

Met Ser Asp Glu Ala Val Glu Glu Arg Arg Ala Leu Ile Lys Arg Lys
115 120 125

Lys Ser Glu Arg Thr Gly Thr Gln Pro Leu Gly Val Gln Gly Leu Thr
130 135 140

Glu Glu Gln Arg Met Met Ile Arg Glu Leu Met Asp Ala Gln Met Lys
145 150 155 160

Thr Phe Asp Thr Thr Phe Ser His Phe Lys Asn Phe Arg Leu Pro Gly
165 170 175

Val Leu Ser Ser Gly Cys Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser
180 185 190

Arg Glu Glu Ala Ala Lys Trp Ser Gln Val Arg Lys Asp Leu Cys Ser
195 200 205

Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly Ser Val Trp Asn
210 215 220

Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu
225 230 235 240

Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser
245 250 255

Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln
260 265 270

Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe
275 280 285

Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu
290 295 300

Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu
305 310 315 320

Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys Leu Gln Leu His
325 330 335

Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu Phe Ser Pro Asp
340 345 350

Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln Leu Gln Glu Gln
355 360 365

Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro
370 375 380

Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met Leu Thr Glu Leu
385 390 395 400

Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu Arg Ile Gln Asp
405 410 415

Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu Phe Gly Ile Thr
420 425 430

Gly Ser

<210> 3

<211> 690

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(690)

<300>

<301> Lehmann et al.

<302> The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions

<303> Journal of Clinical Investigation

<304> 102

<305> 5

<306> 1016-23

<307> 1998

<308> Swiss-Prot/075469; GenBank/AF061056

<309> 1999-01-20

<313> (916)..(1605)

<400> 3

```
gat ctg tgc tct ttg aag gtc tct ctg cag ctg cgg ggg gag gat ggc      48
Asp Leu Cys Ser Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly
1               5                  10                  15

agt gtc tgg aac tac aaa ccc cca gcc gac agt ggc ggg aaa gag atc      96
Ser Val Trp Asn Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile
            20                  25                  30

ttc tcc ctg ctg ccc cac atg gct gac atg tca acc tac atg ttc aaa     144
Phe Ser Leu Leu Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys
        35                  40                  45

ggc atc atc agc ttt gcc aaa gtc atc tcc tac ttc agg gac ttg ccc     192
Gly Ile Ile Ser Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro
    50                  55                  60

atc gag gac cag atc tcc ctg ctg aag ggg gcc gct ttc gag ctg tgt     240
Ile Glu Asp Gln Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys
65                  70                  75                  80

caa ctg aga ttc aac aca gtg ttc aac gcg gag act gga acc tgg gag     288
Gln Leu Arg Phe Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu
                85                  90                  95

tgt ggc cgg ctg tcc tac tgc ttg gaa gac act gca ggt ggc ttc cag     336
Cys Gly Arg Leu Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln
            100                 105                 110

caa ctt cta ctg gag ccc atg ctg aaa ttc cac tac atg ctg aag aag     384
Gln Leu Leu Leu Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys
        115                 120                 125

ctg cag ctg cat gag gag gag tat gtg ctg atg cag gcc atc tcc ctc     432
Leu Gln Leu His Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu
    130                 135                 140

ttc tcc cca gac cgc cca ggt gtg ctg cag cac cgc gtg gtg gac cag     480
Phe Ser Pro Asp Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln
145                 150                 155                 160

ctg cag gag caa ttc gcc att act ctg aag tcc tac att gaa tgc aat     528
Leu Gln Glu Gln Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn
                165                 170                 175

cgg ccc cag cct gct cat agg ttc ttg ttc ctg aag atc atg gct atg     576
Arg Pro Gln Pro Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met
            180                 185                 190

ctc acc gag ctc cgc agc atc aat gct cag cac acc cag cgg ctg ctg     624
Leu Thr Glu Leu Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu
        195                 200                 205

cgc atc cag gac ata cac ccc ttt gct acg ccc ctc atg cag gag ttg     672
Arg Ile Gln Asp Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu
    210                 215                 220

ttc ggc atc aca ggt agc                                             690
Phe Gly Ile Thr Gly Ser
225                 230
```

<210> 4

<211> 230

<212> PRT

<213> Homo sapiens

<400> 4

Asp Leu Cys Ser Leu Lys Val Ser Leu Gln Leu Arg Gly Glu Asp Gly
1 5 10 15

Ser Val Trp Asn Tyr Lys Pro Pro Ala Asp Ser Gly Gly Lys Glu Ile
20 25 30

Phe Ser Leu Leu Pro His Met Ala Asp Met Ser Thr Tyr Met Phe Lys
35 40 45

Gly Ile Ile Ser Phe Ala Lys Val Ile Ser Tyr Phe Arg Asp Leu Pro
50 55 60

Ile Glu Asp Gln Ile Ser Leu Leu Lys Gly Ala Ala Phe Glu Leu Cys
65 70 75 80

Gln Leu Arg Phe Asn Thr Val Phe Asn Ala Glu Thr Gly Thr Trp Glu
85 90 95

Cys Gly Arg Leu Ser Tyr Cys Leu Glu Asp Thr Ala Gly Gly Phe Gln
100 105 110

Gln Leu Leu Leu Glu Pro Met Leu Lys Phe His Tyr Met Leu Lys Lys
115 120 125

Leu Gln Leu His Glu Glu Glu Tyr Val Leu Met Gln Ala Ile Ser Leu
130 135 140

Phe Ser Pro Asp Arg Pro Gly Val Leu Gln His Arg Val Val Asp Gln
145 150 155 160

Leu Gln Glu Gln Phe Ala Ile Thr Leu Lys Ser Tyr Ile Glu Cys Asn
165 170 175

Arg Pro Gln Pro Ala His Arg Phe Leu Phe Leu Lys Ile Met Ala Met
180 185 190

Leu Thr Glu Leu Arg Ser Ile Asn Ala Gln His Thr Gln Arg Leu Leu
195 200 205

Arg Ile Gln Asp Ile His Pro Phe Ala Thr Pro Leu Met Gln Glu Leu

210 215 220

Phe Gly Ile Thr Gly Ser
225 230

<210> 5

<211> 13

<212> PRT

<213> Homo sapiens

<400> 5

Glu Asp Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu Glu
1 5 10

<210> 6

<211> 4604

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (116)..(1399)

<300>

<301> Baker et al.

<302> Cloning and expression of full-length cDNA encoding human vitamin D receptor

<303> Proc. Natl. Acad. Sci. U.S.A.

<304> 85

<305> 10

<306> 3294-98

<307> 1988

<308> Swiss-Prot/11473; GenBank/J03258

<309> 1995-01-14

<313> (116)..(1399)

<400> 6

```
ggaacagctt gtccacccgc cggccggacc agaagccttt gggtctgaag tgtctgtgag      60

acctcacaga agagcacccc tgggctccac ttacctgccc cctgctcctt caggg atg      118
                                                              Met
                                                               1

gag gca atg gcg gcc agc act tcc ctg cct gac cct gga gac ttt gac      166
Glu Ala Met Ala Ala Ser Thr Ser Leu Pro Asp Pro Gly Asp Phe Asp
            5                  10                  15

cgg aac gtg ccc cgg atc tgt ggg gtg tgt gga gac cga gcc act ggc      214
Arg Asn Val Pro Arg Ile Cys Gly Val Cys Gly Asp Arg Ala Thr Gly
        20                  25                  30

ttt cac ttc aat gct atg acc tgt gaa ggc tgc aaa ggc ttc ttc agg      262
Phe His Phe Asn Ala Met Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg
    35                  40                  45

cga agc atg aag cgg aag gca cta ttc acc tgc ccc ttc aac ggg gac      310
Arg Ser Met Lys Arg Lys Ala Leu Phe Thr Cys Pro Phe Asn Gly Asp
50                  55                  60                  65

tgc cgc atc acc aag gac aac cga cgc cac tgc cag gcc tgc cgg ctc      358
Cys Arg Ile Thr Lys Asp Asn Arg Arg His Cys Gln Ala Cys Arg Leu
                70                  75                  80

aaa cgc tgt gtg gac atc ggc atg atg aag gag ttc att ctg aca gat      406
Lys Arg Cys Val Asp Ile Gly Met Met Lys Glu Phe Ile Leu Thr Asp
            85                  90                  95

gag gaa gtg cag agg aag cgg gag atg atc ctg aag cgg aag gag gag      454
Glu Glu Val Gln Arg Lys Arg Glu Met Ile Leu Lys Arg Lys Glu Glu
        100                 105                 110

gag gcc ttg aag gac agt ctg cgg ccc aag ctg tct gag gag cag cag      502
Glu Ala Leu Lys Asp Ser Leu Arg Pro Lys Leu Ser Glu Glu Gln Gln
    115                 120                 125

cgc atc att gcc ata ctg ctg gac gcc cac cat aag acc tac gac ccc      550
Arg Ile Ile Ala Ile Leu Leu Asp Ala His His Lys Thr Tyr Asp Pro
130                 135                 140                 145

acc tac tcc gac ttc tgc cag ttc cgg cct cca gtt cgt gtg aat gat      598
Thr Tyr Ser Asp Phe Cys Gln Phe Arg Pro Pro Val Arg Val Asn Asp
                150                 155                 160

ggt gga ggg agc cat cct tcc agg ccc aac tcc aga cac act ccc agc      646
Gly Gly Gly Ser His Pro Ser Arg Pro Asn Ser Arg His Thr Pro Ser
            165                 170                 175

ttc tct ggg gac tcc tcc tcc tcc tgc tca gat cac tgt atc acc tct      694
Phe Ser Gly Asp Ser Ser Ser Ser Cys Ser Asp His Cys Ile Thr Ser
        180                 185                 190

tca gac atg atg gac tcg tcc agc ttc tcc aat ctg gat ctg agt gaa      742
Ser Asp Met Met Asp Ser Ser Ser Phe Ser Asn Leu Asp Leu Ser Glu
    195                 200                 205

gaa gat tca gat gac cct tct gtg acc cta gag ctg tcc cag ctc tcc      790
Glu Asp Ser Asp Asp Pro Ser Val Thr Leu Glu Leu Ser Gln Leu Ser
210                 215                 220                 225

atg ctg ccc cac ctg gct gac ctg gtc agt tac agc atc caa aag gtc      838
Met Leu Pro His Leu Ala Asp Leu Val Ser Tyr Ser Ile Gln Lys Val
                230                 235                 240
```

```
att ggc ttt gct aag atg ata cca gga ttc aga gac ctc acc tct gag      886
Ile Gly Phe Ala Lys Met Ile Pro Gly Phe Arg Asp Leu Thr Ser Glu
            245                 250                 255

gac cag atc gta ctg ctg aag tca agt gcc att gag gtc atc atg ttg      934
Asp Gln Ile Val Leu Leu Lys Ser Ser Ala Ile Glu Val Ile Met Leu
        260                 265                 270

cgc tcc aat gag tcc ttc acc atg gac gac atg tcc tgg acc tgt ggc      982
Arg Ser Asn Glu Ser Phe Thr Met Asp Asp Met Ser Trp Thr Cys Gly
    275                 280                 285

aac caa gac tac aag tac cgc gtc agt gac gtg acc aaa gcc gga cac     1030
Asn Gln Asp Tyr Lys Tyr Arg Val Ser Asp Val Thr Lys Ala Gly His
290                 295                 300                 305

agc ctg gag ctg att gag ccc ctc atc aag ttc cag gtg gga ctg aag     1078
Ser Leu Glu Leu Ile Glu Pro Leu Ile Lys Phe Gln Val Gly Leu Lys
                310                 315                 320

aag ctg aac ttg cat gag gag gag cat gtc ctg ctc atg gcc atc tgc     1126
Lys Leu Asn Leu His Glu Glu Glu His Val Leu Leu Met Ala Ile Cys
            325                 330                 335

atc gtc tcc cca gat cgt cct ggg gtg cag gac gcc gcg ctg att gag     1174
Ile Val Ser Pro Asp Arg Pro Gly Val Gln Asp Ala Ala Leu Ile Glu
        340                 345                 350

gcc atc cag gac cgc ctg tcc aac aca ctg cag acg tac atc cgc tgc     1222
Ala Ile Gln Asp Arg Leu Ser Asn Thr Leu Gln Thr Tyr Ile Arg Cys
    355                 360                 365

cgc cac ccg ccc ccg ggc agc cac ctg ctc tat gcc aag atg atc cag     1270
Arg His Pro Pro Pro Gly Ser His Leu Leu Tyr Ala Lys Met Ile Gln
370                 375                 380                 385

aag cta gcc gac ctg cgc agc ctc aat gag gag cac tcc aag cag tac     1318
Lys Leu Ala Asp Leu Arg Ser Leu Asn Glu Glu His Ser Lys Gln Tyr
                390                 395                 400

cgc tgc ctc tcc ttc cag cct gag tgc agc atg aag cta acg ccc ctt     1366
Arg Cys Leu Ser Phe Gln Pro Glu Cys Ser Met Lys Leu Thr Pro Leu
            405                 410                 415

gtg ctc gaa gtg ttt ggc aat gag atc tcc tga ctaggacagc ctgtgcggtg   1419
Val Leu Glu Val Phe Gly Asn Glu Ile Ser
        420                 425

cctgggtggg gctgctcctc cagggccacg tgccaggccc ggggctggcg gctactcagc   1479

agccctcctc acccgtctgg ggttcagccc ctcctctgcc acctccccta tccacccagc   1539

ccattctctc tcctgtccaa cctaacccct ttcctgcggg cttttccccg gtcccttgag   1599

acctcagcca tgaggagttg ctgtttgttt gacaaagaaa cccaagtggg ggcagagggc   1659

agaggctgga ggcaggcctt gcccagagat gcctccaccg ctgcctaagt ggctgctgac   1719

tgatgttgag ggaacagaca ggagaaatgc atccattcct cagggacaga gacacctgca   1779

cctcccccca ctgcaggccc cgcttgtcca gcgcctagtg gggtctccct ctcctgcctt   1839

actcacgata aataatcggc ccacagctcc caccccaccc ccttcagtgc ccaccaacat   1899

cccattgccc tggttatatt ctcacgggca gtagctgtgg tgaggtgggt tttcttccca   1959

tcactggagc accaggcacg aacccacctg ctgagagacc caaggaggaa aaacagacaa   2019

aaacagcctc acagaagaat atgacagctg tccctgtcac caagctcaca gttcctcgcc   2079

ctgggtctaa ggggttggtt gaggtggaag ccctccttcc acggatccat gtagcaggac   2139

tgaattgtcc ccagtttgca gaaaagcacc tgccgacctc gtcctccccc tgccagtgcc   2199

ttacctcctg cccaggagag ccagccctcc ctgtcctcct cggatcaccg agagtagccg   2259
```

agagcctgct cccccacccc ctccccaggg gagagggtct ggagaagcag tgagccgcat 2319

cttctccatc tggcagggtg ggatggagga gaagaatttt cagaccccag cggctgagtc 2379

atgatctccc tgccgcctca atgtggttgc aaggccgctg ttcaccacag ggctaagagc 2439

taggctgccg cacccagag tgtgggaagg gagagcgggg cagtctcggg tggctagtca 2499

gagagagtgt ttgggggttc cgtgatgtag ggtaaggtgc cttcttattc tcactccacc 2559

acccaaaagt caaaaggtgc ctgtgaggca ggggcggagt gatacaactt caagtgcatg 2619

ctctctgcag gtcgagccca gcccagctgg tgggaagcgt ctgtccgttt actccaaggt 2679

gggtctttgt gagagtgagc tgtaggtgtg cgggaccggt acagaaaggc gttcttcgag 2739

gtggatcaca gaggcttctt cagatcaatg cttgagtttg gaatcggccg cattccctga 2799

gtcaccagga atgttaaagt cagtgggaac gtgactgccc caactcctgg aagctgtgtc 2859

cttgcacctg catccgtagt tccctgaaaa cccagagagg aatcagactt cacactgcaa 2919

gagccttggt gtccacctgg ccccatgtct ctcagaattc ttcaggtgga aaaacatctg 2979

aaagccacgt tccttactgc agaatagcat atatatcgct taatcttaaa tttattagat 3039

atgagttgtt ttcagactca gactccattt gtattatagt ctaatataca gggtagcagg 3099

taccactgat ttggagatat ttatgggggg agaacttaca ttgtgaaact tctgtacatt 3159

aattattatt gctgttgtta ttttacaagg gtctagggag agacccttgt ttgattttag 3219

ctgcagaact gtattggtcc agcttgctct tcagtgggag aaaaacactt gtaagttgct 3279

aaacgagtca atcccctcat tcaggaaaac tgacagagga gggcgtgact cacccaagcc 3339

atatataact agctagaagt gggccaggac aggccgggcg cggtggctca cgcctgtaat 3399

cccagcagtt tgggaggtcg aggtaggtgg atcacctgag gtcgggagtt cgagaccaac 3459

ctgaccaaca tggagaaacc ctgtctctat taaaaatata aaaaaaaaaa aaaaaaaaaa 3519

tagccgggca tggtggcgca agcctgtaat cccagctact caggaggctg aggcagaaga 3579

attgaaccca ggaggtggag gttgcagtga gctgagatcg tgccgttact ctccaacctg 3639

gacaacaaga gcgaaactcc gtcttagaag tggaccagga caggaccaga ttttggagtc 3699

atggtccggt gtccttttca ctacaccatg tttgagctca gaccccact ctcattcccc 3759

aggtggctga cccagtccct gggggaagcc ctggatttca gaaagagcca agtctggatc 3819

tgggacccct tccttccttc cctggcttgt aactccacca agcccatcag aaggagaagg 3879

aaggagactc acctctgcct caatgtgaat cagaccctac cccaccacga tgtgccctgg 3939

ctgctgggct ctccacctca ggccttggat aatgctgttg cctcatctat aacatgcatt 3999

tgtctttgta atgtcaccac cttcccagct ctcccctctgg ccctgcttct tcggggaact 4059

cctgaaatat cagttactca gccctgggcc ccaccaccta ggccactcct ccaaaggaag 4119

tctaggagct gggaggaaaa gaaaagaggg gaaaatgagt ttttatgggg ctgaacgggg 4179

agaaaaggtc atcatcgatt ctactttaga atgagagtgt gaaatagaca tttgtaaatg 4239

taaaactttt aaggtatatc attataactg aaggagaagg tgccccaaaa tgcaagattt 4299

tccacaagat tcccagagac aggaaaatcc tctggctggc taactggaag catgtaggag 4359

aatccaagcg aggtcaacag agaaggcagg aatgtgtggc agatttagtg aaagctagag 4419

atatggcagc gaaaggatgt aaacagtgcc tgctgaatga tttccaaaga gaaaaaaagt 4479

ttgccagaag tttgtcaagt caaccaatgt agaaagcttt gcttatggta ataaaaatgg 4539

ctcatactta tatagcactt actttgtttg caagtactgc tgtaaataaa tgctttatgc 4599

aaacc 4604

<210> 7

<211> 427

<212> PRT

<213> Homo sapiens

<400> 7

Met Glu Ala Met Ala Ala Ser Thr Ser Leu Pro Asp Pro Gly Asp Phe
1 5 10 15

Asp Arg Asn Val Pro Arg Ile Cys Gly Val Cys Gly Asp Arg Ala Thr
20 25 30

Gly Phe His Phe Asn Ala Met Thr Cys Glu Gly Cys Lys Gly Phe Phe
35 40 45

Arg Arg Ser Met Lys Arg Lys Ala Leu Phe Thr Cys Pro Phe Asn Gly
50 55 60

Asp Cys Arg Ile Thr Lys Asp Asn Arg Arg His Cys Gln Ala Cys Arg
65 70 75 80

Leu Lys Arg Cys Val Asp Ile Gly Met Met Lys Glu Phe Ile Leu Thr
85 90 95

Asp Glu Glu Val Gln Arg Lys Arg Glu Met Ile Leu Lys Arg Lys Glu
100 105 110

Glu Glu Ala Leu Lys Asp Ser Leu Arg Pro Lys Leu Ser Glu Glu Gln
115 120 125

Gln Arg Ile Ile Ala Ile Leu Leu Asp Ala His His Lys Thr Tyr Asp
130 135 140

Pro Thr Tyr Ser Asp Phe Cys Gln Phe Arg Pro Pro Val Arg Val Asn
145 150 155 160

Asp Gly Gly Gly Ser His Pro Ser Arg Pro Asn Ser Arg His Thr Pro

165 170 175

Ser Phe Ser Gly Asp Ser Ser Ser Ser Cys Ser Asp His Cys Ile Thr
180 185 190

Ser Ser Asp Met Met Asp Ser Ser Ser Phe Ser Asn Leu Asp Leu Ser
195 200 205

Glu Glu Asp Ser Asp Asp Pro Ser Val Thr Leu Glu Leu Ser Gln Leu
210 215 220

Ser Met Leu Pro His Leu Ala Asp Leu Val Ser Tyr Ser Ile Gln Lys
225 230 235 240

Val Ile Gly Phe Ala Lys Met Ile Pro Gly Phe Arg Asp Leu Thr Ser
245 250 255

Glu Asp Gln Ile Val Leu Leu Lys Ser Ser Ala Ile Glu Val Ile Met
260 265 270

Leu Arg Ser Asn Glu Ser Phe Thr Met Asp Asp Met Ser Trp Thr Cys
275 280 285

Gly Asn Gln Asp Tyr Lys Tyr Arg Val Ser Asp Val Thr Lys Ala Gly
290 295 300

His Ser Leu Glu Leu Ile Glu Pro Leu Ile Lys Phe Gln Val Gly Leu
305 310 315 320

Lys Lys Leu Asn Leu His Glu Glu Glu His Val Leu Leu Met Ala Ile
325 330 335

Cys Ile Val Ser Pro Asp Arg Pro Gly Val Gln Asp Ala Ala Leu Ile
340 345 350

Glu Ala Ile Gln Asp Arg Leu Ser Asn Thr Leu Gln Thr Tyr Ile Arg
355 360 365

Cys Arg His Pro Pro Pro Gly Ser His Leu Leu Tyr Ala Lys Met Ile
370 375 380

Gln Lys Leu Ala Asp Leu Arg Ser Leu Asn Glu Glu His Ser Lys Gln
385 390 395 400

Tyr Arg Cys Leu Ser Phe Gln Pro Glu Cys Ser Met Lys Leu Thr Pro

405 410 415

Leu Val Leu Glu Val Phe Gly Asn Glu Ile Ser
420 425

<210> 8

<211> 708

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(708)

<300>

<301> Baker et al.

<302> Cloning and expression of full-length cDNA encoding human vitamin D receptor

<303> Proc. Natl. Acad. Sci. U.S.A.

<304> 85

<305> 10

<306> 3294-98

<307> 1988

<308> Swiss-Prot/11473; GenBank/J03258

<309> 1995-01-14

<313> (689)..(1396)

<400> 8

```
acc tct tca gac atg atg gac tcg tcc agc ttc tcc aat ctg gat ctg      48
Thr Ser Ser Asp Met Met Asp Ser Ser Ser Phe Ser Asn Leu Asp Leu
1               5                  10                  15

agt gaa gaa gat tca gat gac cct tct gtg acc cta gag ctg tcc cag      96
Ser Glu Glu Asp Ser Asp Asp Pro Ser Val Thr Leu Glu Leu Ser Gln
            20                  25                  30

ctc tcc atg ctg ccc cac ctg gct gac ctg gtc agt tac agc atc caa     144
Leu Ser Met Leu Pro His Leu Ala Asp Leu Val Ser Tyr Ser Ile Gln
        35                  40                  45
```

```
aag gtc att ggc ttt gct aag atg ata cca gga ttc aga gac ctc acc      192
Lys Val Ile Gly Phe Ala Lys Met Ile Pro Gly Phe Arg Asp Leu Thr
    50                  55                  60

tct gag gac cag atc gta ctg ctg aag tca agt gcc att gag gtc atc      240
Ser Glu Asp Gln Ile Val Leu Leu Lys Ser Ser Ala Ile Glu Val Ile
65                  70                  75                  80

atg ttg cgc tcc aat gag tcc ttc acc atg gac gac atg tcc tgg acc      288
Met Leu Arg Ser Asn Glu Ser Phe Thr Met Asp Asp Met Ser Trp Thr
                85                  90                  95

tgt ggc aac caa gac tac aag tac cgc gtc agt gac gtg acc aaa gcc      336
Cys Gly Asn Gln Asp Tyr Lys Tyr Arg Val Ser Asp Val Thr Lys Ala
            100                 105                 110

gga cac agc ctg gag ctg att gag ccc ctc atc aag ttc cag gtg gga      384
Gly His Ser Leu Glu Leu Ile Glu Pro Leu Ile Lys Phe Gln Val Gly
        115                 120                 125

ctg aag aag ctg aac ttg cat gag gag gag cat gtc ctg ctc atg gcc      432
Leu Lys Lys Leu Asn Leu His Glu Glu Glu His Val Leu Leu Met Ala
    130                 135                 140

atc tgc atc gtc tcc cca gat cgt cct ggg gtg cag gac gcc gcg ctg      480
Ile Cys Ile Val Ser Pro Asp Arg Pro Gly Val Gln Asp Ala Ala Leu
145                 150                 155                 160

att gag gcc atc cag gac cgc ctg tcc aac aca ctg cag acg tac atc      528
Ile Glu Ala Ile Gln Asp Arg Leu Ser Asn Thr Leu Gln Thr Tyr Ile
                165                 170                 175

cgc tgc cgc cac ccg ccc ccg ggc agc cac ctg ctc tat gcc aag atg      576
Arg Cys Arg His Pro Pro Pro Gly Ser His Leu Leu Tyr Ala Lys Met
            180                 185                 190

atc cag aag cta gcc gac ctg cgc agc ctc aat gag gag cac tcc aag      624
Ile Gln Lys Leu Ala Asp Leu Arg Ser Leu Asn Glu Glu His Ser Lys
        195                 200                 205

cag tac cgc tgc ctc tcc ttc cag cct gag tgc agc atg aag cta acg      672
Gln Tyr Arg Cys Leu Ser Phe Gln Pro Glu Cys Ser Met Lys Leu Thr
    210                 215                 220

ccc ctt gtg ctc gaa gtg ttt ggc aat gag atc tcc                      708
Pro Leu Val Leu Glu Val Phe Gly Asn Glu Ile Ser
225                 230                 235
```

<210> 9

<211> 236

<212> PRT

<213> Homo sapiens

<400> 9

```
Thr Ser Ser Asp Met Met Asp Ser Ser Ser Phe Ser Asn Leu Asp Leu
1               5                   10                  15

Ser Glu Glu Asp Ser Asp Asp Pro Ser Val Thr Leu Glu Leu Ser Gln
            20                  25                  30
```

Leu Ser Met Leu Pro His Leu Ala Asp Leu Val Ser Tyr Ser Ile Gln
35 40 45

Lys Val Ile Gly Phe Ala Lys Met Ile Pro Gly Phe Arg Asp Leu Thr
50 55 60

Ser Glu Asp Gln Ile Val Leu Leu Lys Ser Ser Ala Ile Glu Val Ile
65 70 75 80

Met Leu Arg Ser Asn Glu Ser Phe Thr Met Asp Asp Met Ser Trp Thr
85 90 95

Cys Gly Asn Gln Asp Tyr Lys Tyr Arg Val Ser Asp Val Thr Lys Ala
100 105 110

Gly His Ser Leu Glu Leu Ile Glu Pro Leu Ile Lys Phe Gln Val Gly
115 120 125

Leu Lys Lys Leu Asn Leu His Glu Glu Glu His Val Leu Leu Met Ala
130 135 140

Ile Cys Ile Val Ser Pro Asp Arg Pro Gly Val Gln Asp Ala Ala Leu
145 150 155 160

Ile Glu Ala Ile Gln Asp Arg Leu Ser Asn Thr Leu Gln Thr Tyr Ile
165 170 175

Arg Cys Arg His Pro Pro Pro Gly Ser His Leu Leu Tyr Ala Lys Met
180 185 190

Ile Gln Lys Leu Ala Asp Leu Arg Ser Leu Asn Glu Glu His Ser Lys
195 200 205

Gln Tyr Arg Cys Leu Ser Phe Gln Pro Glu Cys Ser Met Lys Leu Thr
210 215 220

Pro Leu Val Leu Glu Val Phe Gly Asn Glu Ile Ser
225 230 235

<210> 10

<211> 951

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(951)

<400> 10

```
atg aaa aaa ggt cat cat cat cat cat cat ggt agt gaa cgg aca ggg       48
Met Lys Lys Gly His His His His His His Gly Ser Glu Arg Thr Gly
1               5                  10                  15

act cag cca ctg gga gtg cag ggg ctg aca gag gag cag cgg atg atg       96
Thr Gln Pro Leu Gly Val Gln Gly Leu Thr Glu Glu Gln Arg Met Met
            20                  25                  30

atc agg gag ctg atg gac gct cag atg aaa acc ttt gac act acc ttc      144
Ile Arg Glu Leu Met Asp Ala Gln Met Lys Thr Phe Asp Thr Thr Phe
        35                  40                  45

tcc cat ttc aag aat ttc cgg ctg cca ggg gtg ctt agc agt ggc tgc      192
Ser His Phe Lys Asn Phe Arg Leu Pro Gly Val Leu Ser Ser Gly Cys
    50                  55                  60

gag ttg cca gag tct ctg cag gcc cca tcg agg gaa gaa gct gcc aag      240
Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser Arg Glu Glu Ala Ala Lys
65                  70                  75                  80

tgg agc cag gtc cgg aaa gat ctg tgc tct ttg aag gtc tct ctg cag      288
Trp Ser Gln Val Arg Lys Asp Leu Cys Ser Leu Lys Val Ser Leu Gln
                85                  90                  95

ctg cgg ggg gag gat ggc agt gtc tgg aac tac aaa ccc cca gcc gac      336
Leu Arg Gly Glu Asp Gly Ser Val Trp Asn Tyr Lys Pro Pro Ala Asp
            100                 105                 110

agt ggc ggg aaa gag atc ttc tcc ctg ctg ccc cac atg gct gac atg      384
Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu Pro His Met Ala Asp Met
        115                 120                 125

tca acc tac atg ttc aaa ggc atc atc agc ttt gcc aaa gtc atc tcc      432
Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser Phe Ala Lys Val Ile Ser
    130                 135                 140

tac ttc agg gac ttg ccc atc gag gac cag atc tcc ctg ctg aag ggg      480
Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln Ile Ser Leu Leu Lys Gly
145                 150                 155                 160

gcc gct ttc gag ctg tgt caa ctg aga ttc aac aca gtg ttc aac gcg      528
Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe Asn Thr Val Phe Asn Ala
                165                 170                 175

gag act gga acc tgg gag tgt ggc cgg ctg tcc tac tgc ttg gaa gac      576
Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu Ser Tyr Cys Leu Glu Asp
            180                 185                 190

act gca ggt ggc ttc cag caa ctt cta ctg gag ccc atg ctg aaa ttc      624
Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu Glu Pro Met Leu Lys Phe
        195                 200                 205

cac tac atg ctg aag aag ctg cag ctg cat gag gag gag gag tat gtg ctg   672
His Tyr Met Leu Lys Lys Leu Gln Leu His Glu Glu Glu Tyr Val Leu
    210                 215                 220

atg cag gcc atc tcc ctc ttc tcc cca gac cgc cca ggt gtg ctg cag      720
Met Gln Ala Ile Ser Leu Phe Ser Pro Asp Arg Pro Gly Val Leu Gln
```

225 230 235 240

cac cgc gtg gtg gac cag ctg cag gag caa ttc gcc att act ctg aag 768
His Arg Val Val Asp Gln Leu Gln Glu Gln Phe Ala Ile Thr Leu Lys
245 250 255

tcc tac att gaa tgc aat cgg ccc cag cct gct cat agg ttc ttg ttc 816
Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro Ala His Arg Phe Leu Phe
260 265 270

ctg aag atc atg gct atg ctc acc gag ctc cgc agc atc aat gct cag 864
Leu Lys Ile Met Ala Met Leu Thr Glu Leu Arg Ser Ile Asn Ala Gln
275 280 285

cac acc cag cgg ctg ctg cgc atc cag gac ata cac ccc ttt gct acg 912
His Thr Gln Arg Leu Leu Arg Ile Gln Asp Ile His Pro Phe Ala Thr
290 295 300

ccc ctc atg cag gag ttg ttc ggc atc aca ggt agc tga 951
Pro Leu Met Gln Glu Leu Phe Gly Ile Thr Gly Ser
305 310 315

<210> 11

<211> 316

<212> PRT

<213> Homo sapiens

<400> 11

Met Lys Lys Gly His His His His His His Gly Ser Glu Arg Thr Gly
1 5 10 15

Thr Gln Pro Leu Gly Val Gln Gly Leu Thr Glu Glu Gln Arg Met Met
20 25 30

Ile Arg Glu Leu Met Asp Ala Gln Met Lys Thr Phe Asp Thr Thr Phe
35 40 45

Ser His Phe Lys Asn Phe Arg Leu Pro Gly Val Leu Ser Ser Gly Cys
50 55 60

Glu Leu Pro Glu Ser Leu Gln Ala Pro Ser Arg Glu Glu Ala Ala Lys
65 70 75 80

Trp Ser Gln Val Arg Lys Asp Leu Cys Ser Leu Lys Val Ser Leu Gln
85 90 95

Leu Arg Gly Glu Asp Gly Ser Val Trp Asn Tyr Lys Pro Pro Ala Asp
100 105 110

Ser Gly Gly Lys Glu Ile Phe Ser Leu Leu Pro His Met Ala Asp Met

115 120 125

Ser Thr Tyr Met Phe Lys Gly Ile Ile Ser Phe Ala Lys Val Ile Ser
130 135 140

Tyr Phe Arg Asp Leu Pro Ile Glu Asp Gln Ile Ser Leu Leu Lys Gly
145 150 155 160

Ala Ala Phe Glu Leu Cys Gln Leu Arg Phe Asn Thr Val Phe Asn Ala
165 170 175

Glu Thr Gly Thr Trp Glu Cys Gly Arg Leu Ser Tyr Cys Leu Glu Asp
180 185 190

Thr Ala Gly Gly Phe Gln Gln Leu Leu Leu Glu Pro Met Leu Lys Phe
195 200 205

His Tyr Met Leu Lys Lys Leu Gln Leu His Glu Glu Glu Tyr Val Leu
210 215 220

Met Gln Ala Ile Ser Leu Phe Ser Pro Asp Arg Pro Gly Val Leu Gln
225 230 235 240

His Arg Val Val Asp Gln Leu Gln Glu Gln Phe Ala Ile Thr Leu Lys
245 250 255

Ser Tyr Ile Glu Cys Asn Arg Pro Gln Pro Ala His Arg Phe Leu Phe
260 265 270

Leu Lys Ile Met Ala Met Leu Thr Glu Leu Arg Ser Ile Asn Ala Gln
275 280 285

His Thr Gln Arg Leu Leu Arg Ile Gln Asp Ile His Pro Phe Ala Thr
290 295 300

Pro Leu Met Gln Glu Leu Phe Gly Ile Thr Gly Ser
305 310 315

<210> 12

<211> 276

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(276)

<300>

<301> Takeshita et al.

<302> Molecular cloning and properties of a full-length putative thyroid hormone receptor coactivator

<303> Endocrinology

<304> 137

<305> 8

<306> 3594-3597

<307> 1996

<308> GenBank/U59302

<309> 1996-08-08

<313> (623)..(710)

<400> 12

```
atg aaa aaa gac agt aaa tac tct caa acc agt cac aaa cta gtg cag      48
Met Lys Lys Asp Ser Lys Tyr Ser Gln Thr Ser His Lys Leu Val Gln
1               5                  10                  15

ctt ttg aca aca act gcc gaa cag cag tta cgg cat gct gat ata gac      96
Leu Leu Thr Thr Thr Ala Glu Gln Gln Leu Arg His Ala Asp Ile Asp
            20                  25                  30

aca agc tgc aaa gat gtc ctg tct tgc aca ggc act tcc aac tct gcc     144
Thr Ser Cys Lys Asp Val Leu Ser Cys Thr Gly Thr Ser Asn Ser Ala
        35                  40                  45

tct gct aac tct tca gga ggt tct tgt ccc tct tct cat agc tca ttg     192
Ser Ala Asn Ser Ser Gly Gly Ser Cys Pro Ser Ser His Ser Ser Leu
    50                  55                  60

aca gaa cgg cat aaa att cta cac cgg ctc tta cag gag ggt agc ccc     240
Thr Glu Arg His Lys Ile Leu His Arg Leu Leu Gln Glu Gly Ser Pro
65                  70                  75                  80

tca gat atc acc act ttg tct gtc gag cct gat taa                     276
Ser Asp Ile Thr Thr Leu Ser Val Glu Pro Asp
                85                  90
```

<210> 13

<211> 91

<212> PRT

<213> Homo sapiens

<400> 13

Met Lys Lys Asp Ser Lys Tyr Ser Gln Thr Ser His Lys Leu Val Gln
1 5 10 15

Leu Leu Thr Thr Thr Ala Glu Gln Gln Leu Arg His Ala Asp Ile Asp
20 25 30

Thr Ser Cys Lys Asp Val Leu Ser Cys Thr Gly Thr Ser Asn Ser Ala
35 40 45

Ser Ala Asn Ser Ser Gly Gly Ser Cys Pro Ser Ser His Ser Ser Leu
50 55 60

Thr Glu Arg His Lys Ile Leu His Arg Leu Leu Gln Glu Gly Ser Pro
65 70 75 80

Ser Asp Ile Thr Thr Leu Ser Val Glu Pro Asp
85 90

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/15701

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : G06F 19/00
US CL : 702/27
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 702/27

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WATKINS et al. The Human Nuclear Xenobiotic Receptor PXR: Structural Determinants of Directed Promiscuity. Science. 22 June 2001. Volume 292, Issue 5525, pages 2329-2333. See entire document.	1-12
X	WATKINS et al. Crystal Structure Of Human Pregame X Receptor Ligand Binding Domain Bound To Sr12813. Protein Data Bank. OCA Atlas for IILH. 08 May 2001. See entire document.	1-12
X	JONES et al. The Pregnane X Receptor: A Promiscuous Xenobiotic Receptor That Has Diveged during Evolution. Molecular Endocrinology. 2000, Volume 14, Number 1, pages 27-36. See especially page 29 line 3.	5

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
10 August 2002 (10.08.2002)

Date of mailing of the international search report
12 SEP 2002

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703)305-3230

Authorized officer
Nikolai M Galitsky, PhD.
Telephone No. (703)308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/15701

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-12

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/15701

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions, which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-12, drawn to a pure PXR ligand binding domain polypeptide in crystalline form.
Group II, claims 13-19, drawn to a method for determining the three-dimensional structure of a crystallized hPXR ligand binding domain polypeptide to a resolution of about 3.0Å or better.
Group III, claim 20, drawn to a method of designing a modulator of a PXR polypeptide.
Group IV, claims 21-24, drawn to a method of designing a modulator that selectively modulates the activity of a human PXR polypeptide.
Group V, claims 25-26, drawn to a method for identifying a PXR modulator by modeling a ligand that fits spatially into a binding cavity or on the surface of the PXR.
Group VI, claims 27-28, drawn to a method of identifying a PXR modulator that selectively modulates the activity of a PXR polypeptide compared to other polypeptides.
Group VII, claims 29-34, and 36-37, drawn to a method of designing a modulator of a PXR polypeptide.
Group VIII, claim 35, drawn to an assay method for identifying a compound that inhibits binding of a ligand to a PXR polypeptide.
Group IX, claims 38-41, drawn to a method of evaluating a candidate therapeutic agent in humans using a mouse model system.
Group X, claims 42-47, drawn to an expression of the mutant mouse PXR polypeptide.
The inventions listed as Groups I-X do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:
The special technical feature for the inventions of Group I is a Pure PXR ligand binding domain polypeptide in crystalline form.
The special technical feature for the inventions of Group II is a method for determining the three-dimensional structure of a crystallized hPXR ligand binding domain polypeptide to a resolution of about 3.0Å.
The special technical feature for the inventions of Group III is a method of designing a modulator of a PXR polypeptide.
The special technical feature for the inventions of Group IV is a method of designing a modulator that selectively modulates the activity of a human PXR polypeptide.
The special technical feature for the inventions of Group V is a method for identifying a PXR modulator by modeling a ligand that fits spatially into a binding cavity or on the surface of the PXR.
The special technical feature for the inventions of Group VI is a method of identifying a PXR modulator that selectively modulates the activity of a PXR polypeptide compared to other polypeptides.
The special technical feature for the inventions of Group VII is a computer-assisted method for identifying a modulator of RNAP activity.
The special technical feature for the inventions of Group XIII is an assay method for identifying a compound that inhibits binding of a ligand to a PXR polypeptide.
The special technical feature for the inventions of Group IX is a method of evaluating a candidate therapeutic agent in humans using a mouse model system.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/15701

The special technical feature for the inventions of Group X is an expression of the mutant mouse PXR polypeptide.

Continuation of B. FIELDS SEARCHED Item 3:
WEST, DB=USPT,PGPB,JPAB,EPAB,DWPI; STN, Non-Patent Literature.
Search termspregnane, receptor, ligand, domain, crystal, polypeptide, xenobiotic, space group.